



LJMU Research Online

Boddy, LM, Noonan, RJ, Rowlands, AV, Hurter, L, Knowles, ZR and Fairclough, SJ

The backwards comparability of wrist worn GENEActiv and waist worn ActiGraph accelerometer estimates of sedentary time in children

<http://researchonline.ljmu.ac.uk/id/eprint/10111/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Boddy, LM, Noonan, RJ, Rowlands, AV, Hurter, L, Knowles, ZR and Fairclough, SJ (2019) The backwards comparability of wrist worn GENEActiv and waist worn ActiGraph accelerometer estimates of sedentary time in children. Journal of Science and Medicine in Sport. ISSN 1440-2440

LJMU has developed [LJMU Research Online](#) for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

1 **Title Page**

2 Title: The backwards comparability of wrist worn GENEActiv and waist worn ActiGraph accelerometer
3 estimates of sedentary time in children

4 **Authors:** Lynne M. Boddy¹, Robert J. Noonan², Alex V. Rowlands^{3,4,5}, Liezel Hurter¹, Zoe R.
5 Knowles¹, Stuart J. Fairclough².

6

7 **Institution and affiliations:**

8 ¹Physical Activity Exchange, Research Institute for Sport and Exercise Sciences, Liverpool John
9 Moores University, Liverpool, UK

10 ²Department of Sport and Physical Activity, Edge Hill University, Ormskirk, UK

11 ³Diabetes Research Centre, University of Leicester, Leicester General Hospital, Leicester, UK;

12 ⁴NIHR Leicester Biomedical Research Centre, UK;

13 ⁵Alliance for Research in Exercise, Nutrition and Activity (ARENA), Sansom Institute for Health
14 Research, Division of Health Sciences, University of South Australia, Adelaide, Australia.

15

16 **Corresponding author:** Dr Lynne M. Boddy, L.M.Boddy@ljmu.ac.uk.

17

18 **Abstract**

19 *Objectives:* To examine the backward comparability of a range of wrist-worn accelerometer estimates
20 of sedentary time (ST) with ActiGraph 100 count·min⁻¹ waist ST estimates.

21 *Design:* Cross-sectional, secondary data analysis

22 *Method:* One hundred and eight 10-11-year-old children (65 girls) wore an ActiGraph GT3X+
23 accelerometer (AG) on their waist and a GENEActiv accelerometer (GA) on their non-dominant wrist
24 for seven days. GA ST data were classified using a range of thresholds from 23-56 mg. ST estimates
25 were compared to AG ST 100 count·min⁻¹ data. Agreement between the AG and GA thresholds was
26 examined using Cronbach's alpha, intraclass correlation coefficients (ICC), limits of agreement (LOA),
27 Kappa values, percent agreement, mean absolute percent error (MAPE) and equivalency analysis.

28 *Results:* Mean AG total ST was 492.4 minutes over the measurement period. Kappa values ranged from
29 0.31-0.39. Percent agreement ranged from 68-69.9%. Cronbach's alpha values ranged from 0.88-0.93.
30 ICCs ranged from 0.59-0.86. LOA were wide for all comparisons. Only the 34 mg threshold produced
31 estimates that were equivalent at the group level to the AG ST 100 count·min⁻¹ data though sensitivity
32 and specificity values of ~64% and ~74% respectively were observed.

33 *Conclusions:* Wrist-based estimates of ST generated using the 34 mg threshold are comparable with
34 those derived from the AG waist mounted 100 count·min⁻¹ threshold at the group level. The 34 mg
35 threshold could be applied to allow group-level comparisons of ST with evidence generated using the
36 ActiGraph 100 count·min⁻¹ method though it is important to consider the observed sensitivity and
37 specificity results when interpreting findings.

38

39 *Keywords:* accelerometry, physical activity, sedentary behaviour, children, raw acceleration signals,
40 measurement

41 **Introduction**

42 Sedentary behaviour (SB) has received increased attention across recent years as a behaviour that may
43 detrimentally affect children's health. Whether SB influences health independent of physical activity
44 (PA) is deemed to be a controversial topic, with some studies demonstrating the negative effects of
45 reallocating moderate-to vigorous PA (MVPA) to SB ^{1,2}, and others reporting limited evidence that SB
46 is associated with health independent of MVPA ³. Nonetheless, researchers are interested in measuring
47 youth movement behaviours including SB to explore health associations, investigate secular trends, and
48 establish intervention effects.

49

50 Sedentary behaviour is defined as any waking behaviour characterised by an energy expenditure of ≤ 1.5
51 metabolic equivalents (METs) while in a sitting, reclining or lying posture ⁴. Despite the SB definition
52 referring to posture, many researchers use accelerometers to classify SB as an absence of, or little,
53 registered dynamic acceleration ⁵. While widely used, it should be noted that this approach does not
54 consider posture.

55

56 Historically children's SB or sedentary time (ST: the time spent for any duration or in any context in
57 sedentary behaviours ⁴) was assessed using waist-worn ActiGraph accelerometers with a threshold of
58 ≤ 100 vertical axis count \cdot min⁻¹ used as the upper boundary for ST. This approach has demonstrated
59 acceptable agreement with measures that classify posture such as the activPAL ⁶ and those that provide
60 an estimate of energy expenditure, for example indirect calorimetry⁷. Recently the field has moved
61 towards that of wrist accelerometry due to superior wear compliance ⁸. Despite observing better
62 compliance, and moderate-to-strong correlations between acceleration data collected using wrist
63 GENEActiv and waist ActiGraph accelerometer placements⁹, wrist placements generally result in
64 higher estimates of physical activity, therefore wrist and hip data are not directly comparable without
65 correcting for these differences ⁸. Researchers have also begun to make use of raw acceleration signal
66 analysis to remove the proprietary nature of counts-based data and improve comparability between
67 different devices ¹⁰. Although this is advantageous for studies moving forward, a wealth of SB data
68 exists using the ≤ 100 count \cdot min⁻¹ threshold applied to data from hip-worn ActiGraphs (for example: ¹¹,

69 ¹). Therefore, the ability to compare new raw acceleration derived estimates with those generated using
70 the ≤ 100 count \cdot min⁻¹ method would be useful for researchers in the field.

71

72 In a previous study authors generated and cross-validated a GENEActiv (GA) wrist threshold of 51 mg
73 with the intention of providing comparable estimates of ST to those generate using the traditional
74 ActiGraph waist-worn ≤ 100 count \cdot min⁻¹ method. The open source R package GGIR was used to
75 calculate average magnitude of dynamic acceleration, known as the Euclidean Norm Minus One
76 (ENMO), from raw acceleration data¹², applying the newly generated (51 mg) and published thresholds
77 for SB^{13,14}. The comparability of ST estimates between the newly generated 51 mg threshold, the other
78 empirical GA raw acceleration thresholds, and ActiGraph waist-worn ≤ 100 count \cdot min⁻¹ data was
79 examined. Results demonstrated a lack of equivalence for the 51 mg threshold and existing GENEActiv
80 empirical wrist ST thresholds^{15, 16}. The study provided some preliminary evidence of group-level
81 agreement for the 36 mg empirical threshold¹³ which was originally intended to classify wrist-worn
82 ActiGraph data. A study limitation, however, was that the individual level agreement was
83 undetermined. At the study conclusion, the authors called for the backwards compatibility of ST
84 estimates to be examined further, both at the individual *and* group levels covering a broad range of
85 ENMO thresholds between the lowest (23 mg) and highest (56 mg) thresholds used to date by
86 researchers in the field. This would enable researchers to establish the most comparable threshold to
87 use when comparing to earlier estimates of ST from ActiGraph data. The lack of evidence related to the
88 comparability of ST estimates currently presents a challenge for researchers when attempting to
89 compare data to those previously recorded using hip and count methods. More investigation is required,
90 therefore, to confirm whether the 36 mg proposed in the previous study represents the optimal threshold
91 to use for this purpose.

92

93 The backwards comparability of wrist generated ENMO assessed MVPA with traditional accelerometer
94 counts-based data using a range of waist-worn ActiGraph MVPA thresholds has been recently
95 demonstrated. The study proposed wrist ENMO thresholds that gave estimates of MVPA that are
96 comparable with waist measured counts-based data classified using empirical ActiGraph thresholds¹⁷.

97 To date, the backwards comparability for the range of ENMO-derived sedentary behaviour/time
98 estimates has not been comprehensively examined. The aim of this secondary data analysis was
99 therefore to extend previous work by using a wide range of both empirical and arbitrary ST thresholds
100 to examine the backwards comparability of wrist-worn accelerometer estimates of ST with waist-worn
101 ActiGraph 100 count·min⁻¹ ST. The study also develops previous work by investigating the extent of
102 backwards compatibility at both the individual and group levels.

103

104 **Methods**

105 This is a secondary data analysis, and the methods for the study have been previously published
106 elsewhere ^{8, 12, 15}. Briefly, after gaining institutional ethical approval, parental/carer consent and child
107 assent, 108 10-11-year-old children were involved in this study (65 girls). Body mass was assessed to
108 the nearest 0.1 kg (Seca, Birmingham, UK) and stature was assessed to the nearest 0.1 cm using a
109 portable height meter (Leicester Height Measure, Seca, Birmingham, UK) during school-based data
110 collection sessions conducted between January - May 2014.

111

112 Two tri-axial accelerometers (GENEActiv; Activinsights, Cambs, Uk and ActiGraph GT3X+;
113 ActiGraph, Pensacola, FL) were used to assess sedentary time. The GENEActiv (GA) was worn on the
114 non-dominant wrist, and the ActiGraph (AG) worn on the right hip for waking hours over seven
115 consecutive days. Participants were instructed to remove the monitors when engaging in water-based
116 activities and also when sleeping. Both monitors were initialised to record at a sampling frequency of
117 100 Hz using the same computer.

118

119 Data generated by the AG devices were processed using ActiLife v 6.11.4 software (ActiGraph,
120 Pensacola, FL). Consistent with previous research ^{15, 18}, for the AG devices, non-wear was defined as
121 20 minutes of consecutive zero counts (1-minute spike tolerance) and was subtracted from daily wear
122 time. A valid day was defined as ≥ 540 min for weekdays ¹⁹ and ≥ 480 min for weekend days ²⁰.
123 Consistent with a previous study ¹⁵, the valid weekend and weekdays with the longest wear for each
124 participant were retained for analysis. Where participants did not have a valid weekend day, only their

125 longest valid weekday was retained for analysis. Data for the included days were converted into 1
126 second csv files, with non-wear times manually removed at a later step in data reduction described later.
127 Sedentary time was defined as $\leq 100 \text{ count}\cdot\text{min}^{-1}$ ²¹ and coded accordingly. Data generated by the GA
128 monitors were saved as binary files after being downloaded using GENEActiv v 2.2 software
129 (Activinsights, Cambs, UK). GA data were processed in R using the GGIR package version 1.1-4 to
130 calculate the ENMO-derived average magnitude of dynamic acceleration. ENMO is vulnerable to
131 calibration errors, therefore to correct for sensor calibration errors, autocalibration was completed²².
132 Non-wear for the GA data was scored using 60 minutes moving windows with 15 minutes increments
133 and imputation was completed²³. ENMO values were expressed in average mg per 1 second epoch²³,
134 and GA data for the corresponding AG week and/or weekend days were retained for further analysis.

135

136 Time stamps for the GA and AG were synchronised, and data were merged resulting in one csv file for
137 each participant. Periods of non-wear were manually removed from both the AG and GA data according
138 to the wear details generated by the ActiLife AG analysis. Therefore all epochs remaining in the dataset
139 contained 'wear' data for both devices. After non-wear periods were removed, data were then reduced
140 to 1-minute epochs and AG data were scored as sedentary or active using vertical axis $100 \text{ count}\cdot\text{min}^{-1}$
141 as the reference value for ST²¹. In the previous study ST thresholds of 23 mg (obtained by solving the
142 Hildebrand et al., (2014) regression equation), 36 mg¹³, 51 mg (newly generated and cross-validated
143 threshold), and 56 mg¹³ were used. In Step 1 of the analysis for the current study we extend these results
144 to examine comparability of GA ST data classified using a wider range of thresholds. This included a
145 recently published ST threshold of 52 mg that was generated by a child-specific calibration circuit²⁴ and
146 arbitrary thresholds of 30 mg, 40 mg and 45 mg which were chosen to cover the range of thresholds.
147 The final thresholds included in Step 1 were therefore: 23 mg, 30 mg, 36 mg, 40 mg, 45 mg, 52 mg and
148 56 mg. This approach resulted in a range of ST thresholds with which to compare to the AG vertical
149 axis $100 \text{ count}\cdot\text{min}^{-1}$ reference.

150

151 Following calculation of descriptive statistics to describe the participant group, the GA ST estimates
152 were compared to the AG $100 \text{ count}\cdot\text{min}^{-1}$ estimates at the group level by calculating Cronbach's alpha,

153 intraclass correlation coefficients (ICC), Kappa values and percent agreement. Individual level
154 estimates were compared by calculating limits of agreement (LOA), correlations between bias and mean
155 sedentary time (AG and GA) and mean absolute percent error (MAPE, %). Null hypothesis testing is
156 not appropriate when considering the comparability between estimates ²⁵, therefore equivalency
157 analysis was also performed to establish the equivalence of group level estimates of ST on average. A
158 95%equivalence test was completed to establish whether the 90% confidence intervals for the range of
159 GA ST thresholds completely fell within the zone of equivalence, defined as $\pm 10\%$ of the mean AG
160 $100 \text{ count}\cdot\text{min}^{-1}$ classified ST.

161 In Step 2 of the analysis, results from step 1 were used to identify the likely range of most comparable
162 thresholds, and further thresholds within this range were then added to the analysis to attempt to find
163 the optimal threshold. Analyses were also completed separately by sex to further examine the
164 comparability of estimates. To provide a more stringent comparison, for the second step in analysis, the
165 zone of equivalence for the group-level equivalence test was reduced to $\pm 5\%$ of the mean AG 100
166 $\text{count}\cdot\text{min}^{-1}$ classified ST. Analysis was completed using IBM SPSS Statistics v.24 (IBM, Armonk, NY)
167 and Microsoft Excel 2016 (Microsoft, Redmond, WA).

168

169 **Results**

170 Participant characteristics, the number of days included in analyses, weekday and weekend day wear
171 times for boys and girls have been published previously for this population ¹⁵, and are displayed in
172 Table 1.

173 [TABLE 1 ABOUT HERE]

174

175

176 Analysis Step 1. Table 2 summarises the comparisons between the $AG \leq 100 \text{ count}\cdot\text{min}^{-1}$ and various
177 GA ST estimates. Kappa values ranged from 0.31-0.39, representing 'fair' agreement ²⁶. Percent
178 agreement ranged from 68-69.9%. Cronbach's alpha values were 0.88 for the 23 mg threshold,
179 suggesting a good level of consistency, where all other Cronbach's alpha values were >0.9 , suggesting
180 excellent levels of consistency. ICCs ranged from 0.59 for the 23 mg threshold (moderate reliability) to
181 0.86 for the 36 mg threshold (good reliability). Supplementary content A displays the Bland-Altman
182 plots for the comparisons. LOA were wide for AG - GA comparisons, with the narrowest limits
183 observed for the AG v 36 mg comparison (lower LOA = -230.47 upper LOA = 194.81 minutes), with
184 systematic bias apparent. All thresholds from 36 mg and above showed negative bias illustrated by
185 mean bias and the correlation results (i.e. higher GA ST estimates than AG). The highest negative bias
186 observed for the 56 mg threshold. MAPE (%) ranged from 15.8% for the 36 mg threshold to 40.7% for
187 the 56 mg. The results of the equivalency analysis found that only the ST estimates generated by the
188 36 mg threshold could be considered statistically equivalent to the $AG \leq 100 \text{ count}\cdot\text{min}^{-1}$ on average at
189 the group level with 90% CI's falling completely within the $\pm 10\%$ zone of equivalence. Thresholds ≤ 30
190 mg appeared to underestimate and ≥ 40 mg appeared to overestimate ST in comparison to the ST
191 reported using $AG \leq 100 \text{ count}\cdot\text{min}^{-1}$. Therefore, for analysis Step 2 thresholds of 34 mg and 35 mg
192 were included to examine the optimum threshold and analyses were repeated for the whole cohort and
193 separately by sex. The zone of equivalence was reduced to $\pm 5\%$ for the group-level equivalency
194 analysis. Table 3 summarises the comparisons between the $AG \leq 100 \text{ count}\cdot\text{min}^{-1}$ and 34 mg and 35 mg
195 GA ST estimates and includes sensitivity and specificity information. Mean bias was low for the 34 mg
196 threshold though wide limits of agreement were observed for both thresholds and MAPE% was similar
197 to that observed for the 36 mg threshold at 16.2% (34 mg) and 15.8% (35 mg). Sensitivity values (true
198 positive) were similar between the thresholds, at 63.6% and 64.8% for 34 mg and 35 mg respectively.
199 Specificity values (true negative) were also similar, at 74.2% and 73.4% for the 34 mg and 35 mg
200 thresholds respectively. Boys' data displayed wider limits of agreement, higher MAPE% and slightly

201 higher sensitivity values for both thresholds in comparison to girls, though % agreement, Cronbach's
202 alpha, ICC, Kappa and and specificity values were similar. The results of the equivalence analysis for
203 all threshold comparisons are displayed in Figure 1. Only the ST estimates generated by the 34 mg
204 threshold were statistically equivalent to the AG ≤ 100 count·min⁻¹ on average at the group level with
205 90% CI's falling completely within the $\pm 5\%$ zone of equivalence.

206

207 [TABLE 2 ABOUT HERE]

208

209 [TABLE 3 ABOUT HERE]

210

211 [FIGURE 1 ABOUT HERE]

212 Discussion

213 A wealth of existing accelerometer data has used the threshold of ≤ 100 count \cdot min $^{-1}$ applied to waist-
214 worn ActiGraphs to determine time spent sedentary (for example large studies using the International
215 Children's Accelerometry Database ^{11, 1}). As the discipline moves increasingly towards raw
216 acceleration data processing and wrist-worn monitors, the ability for researchers to compare data
217 between studies that have used counts-based processing methods and waist-worn monitors is important.
218 The aim of this study was therefore to examine the backwards comparability of wrist-worn
219 accelerometer estimates of sedentary time (ST) with ActiGraph 100 count \cdot min $^{-1}$ waist ST estimates
220 using a range of empirically determined and arbitrary raw acceleration thresholds for wrist-worn
221 monitors.

222 This study has demonstrated moderate to excellent ICC values, and moderate to good Cronbach's alpha
223 values at the group level for all the GA thresholds. Despite these results, Kappa values were 'fair' and
224 large MAPE values (individual level) were observed. In addition, wide limits of agreement (individual
225 level) were observed between all GA thresholds and the AG standard. Systematic bias was evident,
226 indicating that as estimates of ST increased so did the bias. Equivalency analysis found no thresholds
227 produced estimates of ST that could be considered statistically equivalent on average at the group level
228 in comparison to the AG standard with the exception of the 34 mg threshold. The wide limits of
229 agreement, MAPE and bias results, in the presence of high consistency, as evidence by high ICC and
230 Cronbach's alpha, highlights the importance of considering a range of analyses at the individual and
231 group levels when examining the comparability of ST estimates.

232 Our previous work called for studies to investigate the backwards compatibility of ST estimates ¹⁵. In
233 the present study this was addressed by examining a broad range of thresholds, both empirically
234 determined and arbitrary, and out of the selected thresholds the 34 mg threshold provided the ST
235 estimates most comparable to the ActiGraph 100 count \cdot min $^{-1}$ waist ST at the group level. Despite this,
236 the limits of agreement showed that ST estimates generate using the 34 mg threshold ranged from \pm ~
237 3 hrs in comparison to the AG estimates, therefore suggesting the 34 mg should not be used for
238 individual level comparison. In our previous study we established that the 36 mg threshold, provided

239 equivalent estimates of wrist ST for the GENEActiv as the ≤ 100 count \cdot min⁻¹ standard at the group level
240 when using a $\pm 10\%$ zone of equivalence. The present study that used a more stringent $\pm 5\%$ zone of
241 equivalence suggests that 34 mg may provide a more accurate comparison at the group level, and
242 furthermore suggests that lower and higher thresholds across the range are not appropriate for this
243 purpose. ActiGraph accelerometers are known to produce lower ENMO values than GA devices ⁸,
244 though recent evidence suggests the GA and AG devices provide equivalent estimates between the 30-
245 50 mg range ¹⁶. Irrespective of potential differences between devices, at the group level the 34 mg
246 threshold provided the most comparable estimates of ST to the AG hip ≤ 100 count \cdot min⁻¹ standard, so
247 could be used for comparative purposes across studies moving forward.

248 Despite exhibiting group level equivalency, the sensitivity and specificity values suggest that for every
249 100 minutes of ST classified by the ActiGraph, the GA 34 mg threshold would classify ~64 minutes of
250 ST. Therefore any comparisons between studies using the wrist worn 34 mg threshold and studies using
251 the waist worn AG ≤ 100 count \cdot min⁻¹ method should bear the sensitivity and specificity results in mind
252 when interpreting findings. Furthermore, the 5% zone of equivalence provides a range of ~50 minutes
253 of sedentary time which the 34 mg estimates fell within. Whether a potential difference of ± 50 minutes
254 is clinically meaningful or whether that would provide estimates that are sensitive to change is open to
255 debate. Recent evidence suggests that the reallocation of 15 minutes of sedentary time to moderate to
256 vigorous physical activity (MVPA) predicted changes in obesity and fitness outcomes in children².
257 However, such evidence relies on sedentary time to be reallocated to MVPA, and the impact of
258 reallocation of time to light intensity physical activity or stationary behaviours independent of MVPA
259 remains unclear. Where group-level comparisons with data collected using the AG hip ≤ 100 count \cdot min⁻¹
260 standard are useful the 34 mg threshold can be applied, though where precise estimates of sedentary
261 time or behaviour are required to demonstrate intervention effectiveness or individual level changes
262 alternative methods may be required. Whether the estimates of ST from the GA and AG reflect actual
263 ST remains open to debate. Indeed it is questionable whether the absence or low levels of acceleration
264 should be used in isolation to classify ST, especially considering the postural component that is integral
265 to the definition of sedentary behaviour. Examining the accuracy of measuring ST was not the aim of

266 the present study *per se*, and as such represents a different research question to be addressed in the
267 future. There are, however, ways of processing accelerometer data to classify posture that do not require
268 the use of additional devices or monitoring periods. One example is the sedentary sphere ²⁷, which
269 classifies assumed postural changes based on acceleration signals, arm orientation and wrist orientation.
270 Although this approach has shown promise in adult populations, it has not yet been validated in children
271 and so its utility in this population has not been established. Nonetheless, estimates based on new
272 approaches, irrespective of the method, still raises questions regarding the comparability with the large
273 volume of existing literature therefore a pragmatic solution is warranted.

274 There are some limitations to the present study. We used a 1-minute epoch to determine time spent
275 sedentary to allow a comparison to the AG hip ≤ 100 count·min⁻¹ standard. The majority of children's
276 ST data using AG hip ≤ 100 count·min⁻¹ utilises 1-minute epochs, therefore this approach was necessary
277 to address the study aims. It is well established that children's physical activity behaviours are sporadic
278 in nature ^{28, 29} and though high frequency monitoring is required to detect movement at higher
279 intensities, the 1 minute epoch is unlikely to influence recorded ST which is generally accrued in bouts
280 lasting >2 minutes ³⁰. In addition, the group of participants involved in this study were all from the same
281 geographical location in North-West England and a narrow age-range, therefore their ST behaviours
282 may not be representative of different populations and groups. We included a maximum of 2 days of
283 data (one weekend and one weekday) for each participant, therefore the sedentary levels of participants
284 are not reflective of their habitual patterns. However, the volume of data included allows for comparison
285 between devices and signal classification, where 7 day's data would have been prohibitive in terms of
286 file size. Furthermore, a waking hours accelerometer protocol was used. Therefore recent studies using
287 24-hour protocols may require further investigation to examine the backward compatibility of data,
288 including sleep classification in addition to ST estimates.

289 To the best of our knowledge, this is the first study to examine the backward comparability of wrist
290 assessed sedentary time with ActiGraph 100 count·min⁻¹ waist ST estimates. The results of the study
291 suggest that the 34 mg threshold produced the most comparable estimates of ST and could be used to

292 classify data for group-level comparison with previously published studies that used the 100 count·min⁻¹
293 threshold.

294

295 **Conclusions**

296 Despite observing high ICC and Cronbach's alpha values, the results suggest that the all but one of the
297 wrist mounted, raw acceleration derived ST estimates should not be directly compared with those
298 derived from the 100 count·min⁻¹ waist mounted AG threshold. The 34 mg threshold may provide
299 comparable ST estimates at the group level, and future studies could use the 34 mg threshold when
300 comparing ENMO derived ST estimates group level estimates previously published using the 100
301 count·min⁻¹ approach though it is important to consider the observed sensitivity and specificity results
302 when interpreting findings.

303

304 **Practical Implications**

- 305 • Many previous studies estimated children's sedentary time using waist-mounted ActiGraph
306 accelerometers and the 100 count·min⁻¹ threshold.
- 307 • The backward comparability of wrist-worn raw acceleration derived sedentary time estimates with
308 the wealth of data collected using waist-mounted ActiGraphs is unknown.
- 309 • This study found that the 34 mg threshold could be applied to wrist accelerometer data to provide
310 estimates of sedentary time that are equivalent to the ActiGraph waist-worn 100 count·min⁻¹ on
311 average at the group level, though the sensitivity and specificity values observed in this study should
312 be considered when interpreting findings.

313 **Acknowledgements**

314 The authors would like to thank the children and schools that were involved this research. AR is
315 with the National Institute for Health Research (NIHR) Biomedical Research Centre based at
316 University Hospitals of Leicester and Loughborough University, the National Institute for Health
317 Research Collaboration for Leadership in Applied Health Research and Care – East Midlands
318 (NIHR CLAHRC – EM) and the Leicester Clinical Trials Unit. The views expressed are those of
319 the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

320

- 322 1. Hansen BH, Anderssen SA, Andersen LB, et al. Cross-Sectional Associations of Reallocating
323 Time Between Sedentary and Active Behaviours on Cardiometabolic Risk Factors in Young
324 People: An International Children's Accelerometry Database (ICAD) Analysis. *Sports Med.*
325 2018; 48(10):2401-2412.
- 326 2. Fairclough SJ, Dumuid D, Mackintosh KA, et al. Adiposity, fitness, health-related quality of
327 life and the reallocation of time between children's school day behaviours: A compositional
328 data analysis. *Preventive Medicine Reports.* 2018; 11:254-261.
- 329 3. Cliff DP, Hesketh KD, Vella SA, et al. Objectively measured sedentary behaviour and health
330 and development in children and adolescents: systematic review and meta-analysis. *Obes*
331 *Rev.* 2016; 17(4):330-344.
- 332 4. Tremblay MS, Aubert S, Barnes JD, et al. Sedentary behavior network (SBRN)- Terminology
333 Consensus Project process and outcome. *International Journal of Behavioural Nutrition and*
334 *Physical Activity.* 2017; 14(75).
- 335 5. Lubans DR, Hesketh K, Cliff DP, et al. A systematic review of the validity and reliability of
336 sedentary behaviour measures used with children and adolescents. *Obes Rev.* 2011;
337 12(10):781-799.
- 338 6. Ridgers ND, Salmon J, Ridley K, E. OC, Arundell L, Timperio A. Agreement between activPAL
339 and ActiGraph for assessing children's sedentary time. *Int J Behav Nutr Phys Act.* 2012;
340 9(1):15.
- 341 7. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective
342 measures of physical activity for children. *J Sports Sci.* 2008; 26(14):1557-1565.
- 343 8. Fairclough SJ, Noonan R, Rowlands AV, Van Hees V, Knowles Z, Boddy LM. Wear Compliance
344 and Activity in Children Wearing Wrist- and Hip-Mounted Accelerometers. *Med Sci Sports*
345 *Exerc.* 2016; 48(2):245-253.
- 346 9. Rowlands AV, Frayssse F, Catt M, et al. Comparability of measured acceleration from
347 accelerometry-based activity monitors. *Med Sci Sports Exerc.* 2015; 47(1):201-210.
- 348 10. Rowlands AV, Olds TS, Hillsdon M, et al. Assessing sedentary behavior with the GENEActiv:
349 introducing the sedentary sphere. *Med Sci Sports Exerc.* 2014; 46(6):1235-1247.
- 350 11. Cooper AR, Goodman A, Page AS, et al. Objectively measured physical activity and sedentary
351 time in youth: the International children's accelerometry database (ICAD). *Int J Behav Nutr*
352 *Phys Act.* 2015; 12:113.
- 353 12. Noonan RJ, Boddy LM, Kim Y, Knowles ZR, Fairclough SJ. Comparison of children's free-living
354 physical activity derived from wrist and hip raw accelerations during the segmented week. *J*
355 *Sports Sci.* 2017; 35(21):2067-2072.
- 356 13. Hildebrand M, Hansen BH, van Hees VT, Ekelund U. Evaluation of raw acceleration sedentary
357 thresholds in children and adults. *Scand J Med Sci Sports.* 2016.
- 358 14. Hildebrand M, VANH, Hansen BH, Ekelund U. Age group comparability of raw
359 accelerometer output from wrist- and hip-worn monitors. *Med Sci Sports Exerc.* 2014;
360 46(9):1816-1824.
- 361 15. Boddy LM, Noonan RJ, Kim Y, et al. Comparability of children's sedentary time estimates
362 derived from wrist worn GENEActiv and hip worn ActiGraph accelerometer thresholds.
363 *Journal of science and medicine in sport / Sports Medicine Australia.* 2018.
- 364 16. Rowlands AV, Mirkes EM, Yates T, et al. Accelerometer-assessed Physical Activity in
365 Epidemiology: Are Monitors Equivalent? *Med Sci Sports Exerc.* 2018; 50(2):257-265.
- 366 17. Rowlands AV, Cliff DP, Fairclough SJ, et al. Moving Forward with Backward Compatibility:
367 Translating Wrist Accelerometer Data. *Med Sci Sports Exerc.* 2016; 48(11):2142-2149.
- 368 18. Catellier DJ, Hannan PJ, Murray DM, et al. Imputation of missing data when measuring
369 activity by accelerometry. *Med Sci Sports Exerc.* 2005; 37(Suppl 11):S555-S562.

- 370 19. Rich C, Geraci M, Griffiths L, Sera F, Dezateux C, Cortina-Borja M. Quality control methods in
371 accelerometer data processing: defining minimum wear time. *PLoS one*. 2013; 8(6):e67206.
- 372 20. Rowlands AV, Pilgrim EL, Eston RG. Patterns of habitual activity across weekdays and
373 weekend days in 9-11-year-old children. *Prev Med*. 2008; 46(4):317-324.
- 374 21. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for
375 predicting activity intensity in youth. *Med Sci Sports Exerc*. 2011; 43(7):1360-1368.
- 376 22. van Hees VT, Fang Z, Langford J, et al. Autocalibration of accelerometer data for free-living
377 physical activity assessment using local gravity and temperature: an evaluation on four
378 continents. *Journal of applied physiology*. 2014; 117(7):738-744.
- 379 23. van Hees VT, Gorzelniak L, Dean Leon EC, et al. Separating movement and gravity
380 components in an acceleration signal and implications for the assessment of human daily
381 physical activity. *PLoS One*. 2013; 8(4):e61691.
- 382 24. Hurter L, Fairclough SJ, Knowles ZR, Porcellato LA, Cooper-Ryan AM, Boddy LM. Establishing
383 Raw Acceleration Thresholds to Classify Sedentary and Stationary Behaviour in Children.
384 *Children (Basel)*. 2018; 5(12).
- 385 25. Dixon PM, Saint-Maurice PF, Kim Y, Hibbing P, Bai Y, Welk GJ. A primer on the use of
386 equivalence testing for evaluating measurement agreement. *Medicine and Science in Sports
387 and Exercise*. 2017; Ahead of Print.
- 388 26. Altman D. *Practical Statistics for Medical Research*, London, Chapman and Hall; 1991.
- 389 27. Rowlands AV, Yates T, Olds TS, Davies M, Khunti K, Edwardson CL. Sedentary Sphere: Wrist-
390 Worn Accelerometer-Brand Independent Posture Classification. *Med Sci Sports Exerc*. 2016;
391 48(4):748-754.
- 392 28. Bailey RC, Olson J, Pepper SL, Porszasz J, Barstow TJ, Cooper DM. The level and tempo of
393 children's physical activities: an observational study. *Med Sci Sports Exerc*. 1995; 27(7):1033-
394 1041.
- 395 29. Baquet G, Stratton G, Van Praagh E, Berthoin S. Improving physical activity assessment in
396 prepubertal children with high-frequency accelerometry monitoring: a methodological issue.
397 *Prev Med*. 2007; 44(2):143-147.
- 398 30. Verloigne M, Ridgers ND, Chinapaw M, et al. Patterns of objectively measured sedentary
399 time in 10- to 12-year-old Belgian children: an observational study within the ENERGY-
400 project. *BMC Pediatr*. 2017; 17(1):147.

401

402

403 Table and Figure Legends

404 Table 1. Mean (SD) anthropometric, wear time and number of days included within analysis for boys

405 and girls

	Boys N = 43		Girls N = 65	
	Mean or	SD	Mean or	SD
	Frequency		Frequency	
Age (y)	10	0.4	10	0.3
Height (cm)	139.5	7.9	138	7.4
Body mass (kg)	35.6	8.2	34.2	8.6
BMI (kg·m ⁻²)	18.2	3.00	17.8	3.2
ActiGraph weekday wear (min·day ⁻¹)	739.9	115.6	738.8	100.4
ActiGraph weekend day wear (min·day ⁻¹)	631.8	110.8	661.5	108.3
ActiGraph valid weekdays included	41	N/A	64	N/A
ActiGraph valid weekend days included	30	N/A	46	N/A
Total valid included days	71	N/A	110	N/A

406

407

408

409

410 Table 2. Comparisons between the ActiGraph ≤ 100 count·min⁻¹ standard and GA thresholds

411

Criterion	GA threshold	Sedentary time (mins/included days)	Mean Bias (mins)	Bias vs Mean ST correlation	Cronbach's Alpha	ICC (Single measures)	Limits of agreement Lower	Limits of Agreement upper	Mean absolute percent error % (SD)	% agreement	Kappa
ActiGraph ≤ 100 count·min ⁻¹		492.4									
	23mg	342.6	149.8	-.11	0.88	0.59	-88.38	387.94	35.5 (18.2)	68	0.31
	30mg	440.4	52	-.33**	0.91	0.82	-166.41	270.45	20.2 (16.5)	69.4	0.36
	36mg	510.2	-17.8	-.49**	0.93	0.86	-230.47	194.81	15.8 (15.7)	69.8	0.38
	40mg	554.8	-62.4	-.58**	0.93	0.83	-276.22	151.36	18.2 (16.3)	69.9	0.39
	45mg	603.6	-111.2	-.67**	0.93	0.77	-327.29	104.95	25.1 (16.4)	69.5	0.39
	52mg	660	-167.6	-.75**	0.93	0.69	-391.09	55.91	34.9 (16.7)	68.8	0.39
	56mg	692.3	-199.9	-.79**	0.93	0.63	-430.48	30.7	40.7 (18)	68.2	0.38

412

414 Table 3. Comparisons between the ActiGraph ≤ 100 count·min⁻¹ standard and 34 mg and 35 mg GA thresholds.

415

Criterion	GA threshold	Sedentary time (mins)	Mean Bias (mins)	Bias vs Mean ST correlation	Cronbach's Alpha	ICC (Single measures)	Limits of agreement Lower	Limits of Agreement upper	Mean absolute percent error % (SD)	% agreement	Kappa	Sensitivity %	Specificity %
ActiGraph ≤ 100 count·min ⁻¹		492.4											
	34 mg	489.6	2.8	-.38**	0.92	0.86	-219.51	216.18	16.2 (15.8)	69.8	0.38	63.6	74.2
	35 mg	501	-8.6	-.47**	0.93	0.86	-220.88	203.71	15.8 (15.7)	69.8	0.38	64.8	73.4
Boys													
ActiGraph ≤ 100 count·min ⁻¹		499.7											
	34 mg	494.7	5	-.47**	0.93	0.87	-236.3	246.3	19 (16.8)	70.4	0.40	65.6	74.2
	35 mg	505.5	-5.8	-.52**	0.93	0.87	-247.3	235.6	19 (16.2)	70.4	0.40	66.7	73.3
Girls													
ActiGraph ≤ 100 count·min ⁻¹		487.6											
	34 mg	486.1	1.4	-.31**	0.92	0.85	-193.2	196	14.4 (14.9)	69.4	0.37	62.3	74.2
	35 mg	498	-10.4	-.42**	0.92	0.86	-202.9	182.1	13.7 (15)	69.4	0.37	63.6	73.5

416 Figure 1. ActiGraph ≤ 100 count \cdot min $^{-1}$ $\pm 5\%$ zone of equivalence (467.7 minutes- 517 minutes, dotted
417 lines) and 90% confidence intervals for the GENEActiv sedentary time estimates classified using nine
418 thresholds

