CHAPTER 5

Measuring physical activity in juvenile idiopathic arthritis: activity diary versus accelerometer

G.J.F. Joyce Bos*
Wineke Armbrust*
Jan H.B. Geertzen
Pieter J.J. Sauer
Pieter U. Dijkstra
Otto T.H.M. Lelieveld

*Joint first co-author

The Journal of Rheumatology 2017;44(8):1249-1256
ABSTRACT

Objective
(1) To determine convergent validity of an activity diary (AD) and accelerometer (Actical brand/Phillips-Respironics) in measuring physical activity (PA) in children with juvenile idiopathic arthritis (JIA). (2) To determine how many days give reliable results. (3) To analyze effects of correcting accelerometer data for non-wear.

Methods
Children with JIA (8-13 years) were recruited from 3 Dutch pediatric rheumatology centers. Physical activity was assessed for 7 days with an AD and accelerometer, and was expressed as mean minutes/day of rest, light PA (LPA), moderate to vigorous PA (MVPA), and PA level (PAL). To analyze convergent validity, intraclass correlation coefficients (ICC) were calculated and paired sample Student t tests were performed. The required number of days to achieve reliable results was calculated using the Spearman-Brown prophecy formula.

Results
Convergent validity between AD and accelerometer was moderate for rest and PAL (ICC 0.41). ICC for LPA and MVPA were < 0.24. AD overestimated PAL and MVPA compared with the accelerometer. Wearing the accelerometer for 7-19 days gave reliable PA estimates on group and individual levels. For the AD, 13-36 days were needed. Adjusting accelerometer data for non-wear resulted in a clinically relevant higher mean number of minutes/day spent in LPA (effect size 1.12), but not in MVPA (effect size 0.44).

Conclusion
Convergent validity between AD and accelerometer is moderate to poor. In children with JIA, 1-week assessment with an accelerometer is sufficient to measure PA (all levels) reliably. On an individual level and for clinical use, 3 weeks are required. Additional use of an AD enables correction for non-wear of accelerometer data.
INTRODUCTION

Physical activity (PA), increasing energy expenditure above basal metabolic level\(^1\), contributes to prevention of several chronic conditions, improves psychological health, and is associated with longevity and prevention of all-cause mortality\(^2\)–\(^5\). For patients with juvenile idiopathic arthritis (JIA), it is equally important to profit from these benefits, because evidence shows PA is safe and does not damage joints\(^6\)–\(^8\). PA is reduced in children and adolescents with JIA\(^9\)–\(^16\).

PA can be expressed as total energy expenditure (TEE) in kilojoules (kJ) or kilocalories per day, where TEE is the sum of the basal metabolic rate, diet-induced thermogenesis, and activity-related energy expenditure (AEE). Another way to express PA is by PA level [PAL; TEE (kJ) divided by the basal metabolic rate (kJ)]. PA can be categorized into rest, light PA (LPA), and moderate to vigorous PA (MVPA)\(^17\),\(^18\).

To determine PA, many methods can be used\(^18\)–\(^20\). In JIA, self and proxy reports, questionnaires, recall diaries, and accelerometers, or combinations of these methods have been used to determine PA\(^9\)–\(^15\),\(^21\)–\(^23\).

Because PA can vary from day to day, increasing the number of days measured will improve reliability of measurements but will increase the burden for patients and may decrease adherence. The number of days in which PA was measured previously ranged from 1 to 7\(^9\)–\(^12\),\(^14\),\(^15\),\(^21\)–\(^23\). Further, the number of days needed to determine PA reliably depends on the type of instrument used and on patient characteristics\(^24\).

In general, questionnaires and recall diaries tend to overestimate PA\(^25\),\(^26\). However, accelerometers underestimate PA, while they do not or insufficiently record certain types of activity, in particular, nonambulatory PA with arm and or leg movements\(^27\). The underestimation is enforced by non-wear during activities such as swimming. Therefore, it has been suggested to combine 2 or more techniques to improve the accuracy of measurements\(^20\). In a study in 13- and 15-year-old adolescents, PA was determined using an accelerometer and an activity diary (AD) to register activities while the accelerometer was not worn. Significantly higher levels of MVPA were found when the results were corrected for non-wear\(^28\).
The aims of our study were to (1) determine convergent validity of a 7-day AD and an accelerometer in children with JIA, (2) determine how many days of PA needed to be assessed to obtain reliable results using an AD and accelerometer, and (3) analyze the effect of combining the 2 instruments by using the AD to correct for non-wear of the accelerometer.

**PATIENTS AND METHODS**

**Study design**

Our current study covers a cross-sectional design of a 7-day observation period.

**Subjects**

Participants were children with JIA, aged 8 to 13 years, participating in the Rheumates@Work study, a multicenter trial to evaluate the effects of an internet-based cognitive behavioural program on PA levels (trial number SRCTN92733069)\(^{29}\). For our study, baseline measurements of Rheumates@Work were used. Children were recruited from 3 pediatric rheumatology outpatient clinics in The Netherlands: the Beatrix Children's Hospital of the University Medical Center Groningen; the Wilhelmina Children's Hospital of the University Medical Center Utrecht; and from Reade, Center for Rehabilitation, Amsterdam, from January 2011 until September 2012. The medical ethics research board of all 3 centers approved the study (NL34044.042.10). All patients with JIA diagnosed according to the international league of associations for rheumatology criteria\(^{30}\) were asked to join in the Rheumates@Work study. Children willing to participate filled in an informed consent form and were invited to participate at their own clinic, where disease activity was scored and comorbidity was registered by a pediatric rheumatologist. On the same day, the accelerometer and AD were handed out. The children and one of their parents were verbally and in writing instructed on how to wear and use the accelerometer and AD, simultaneously. Inclusion criteria for our study were disease activity lower than 2 centimeters on a physician’s global assessment scale (0-10). Exclusion criteria were comorbidity that affected maximum exercise capacity and PA, and insufficient proficiency of the Dutch language. Patients without a completed AD and or accelerometer for 7 consecutive days were also excluded from our study.
Accelerometry

An Actical accelerometer (Phillips-Respironics) was worn with an elastic belt over the right hip near the anterior superior iliac spine. This accelerometer has been validated for children aged 7 years up to 18 years of age (sensitivity 86%-97% and specificity 66%-80%)\textsuperscript{31}. It contains an omnidirectional accelerometer that measures occurrence and intensity of motion. This information was used to calculate activity counts per time unit (60 seconds in our study) and AEE in kilocalories per day. TEE (MJ) was calculated with the formula \((\text{AEE} \times 4.1868 ÷ 1000 + \text{BMR}) ÷ 0.9\), where BMR is the basal metabolic rate\textsuperscript{22}. Data from the accelerometer were stored in an Excel file as counts per minute, giving 1440 timepoints per day. Higher counts per minute correspond with higher PA intensity. Cutoff points were used for rest, LPA, and MVPA\textsuperscript{31}. Accelerometer data were visually inspected with help of an actogram, a graphic representation of activity counts per minute, and non-wear time was observed and compared with non-wear time in the Excel file. Non-wear time was defined as 60 consecutive minutes of 0 counts, with allowance for 1 or 2 minutes of counts between 0 and 100. Accelerometer measurements were considered valid when the wearing time summed 6 hours on weekends or 8 hours on weekdays.

Activity Diary

The AD is a reliable instrument for measuring PA in children from 10 years of age and up (intraclass correlations 0.86-0.95)\textsuperscript{32}. It was validated in 15-year-olds using the doubly labeled water method (gold standard for measuring PA), showing a mean difference of 0.01 in PAL and with limits of agreement between -0.47 and 0.49\textsuperscript{33}. Every quarter of an hour, the dominant activity was scored with a number 1-9 (Appendix 1). In case children or parents where in doubt about giving the correct number for the activity, children or parents could contact the investigator or could describe the type of activity in de AD. In cases where a 15-minutes period had more than 1 entry, the first or second entry was chosen alternately.

In case of missing values, children and parents were asked to recall the activity for that period. If there were still missing values present, missing values were corrected to enable the calculation of PAL. Missing values between 9 p.m. and 7 a.m. were imputed with a 1, because this was considered to be sleeping time. Children were instructed to draw a smiley face in the AD at the time the accelerometer was put on in the morning and when it was taken off in the evening. When children forgot to give a number for their activity and only drew a smiley face to
indicate that the accelerometer was worn, missing values were imputed by the activity of the prior 15 minutes. In case of missing data and children had drawn a smiley face indicating that the accelerometer was taken off, data were imputed with the activity of the next 15 minutes. When 4 or fewer missing values remained, they were substituted by activity 2. In cases of more than 4 remaining missing values, the AD was excluded from the analysis.

PA was expressed as PAL and time (minutes) spent at rest, LPA, and MVPA. Corresponding energy expenditure was calculated with known formulas (Appendix 1)\textsuperscript{25,33,34}. To calculate TEE, the energy cost of all 15-minutes periods were summed and divided by 96. PAL was calculated by dividing the TEE for each day with the BMR\textsuperscript{22,33}.

**Correction of accelerometer data for non-wear**

Rest as measured with the accelerometer was compared to AD data. When LPA or MVPA was reported in the AD and the accelerometer data showed rest, we assumed non-wear. Non-wear was corrected by adding up the number of minutes of LPA or MVPA, reported in the AD, and subtracting the equivalent number of minutes from the total minutes spent at rest. No corrections were made for PAL, because algorithms to calculate energy expenditure use activity counts for each individual minute\textsuperscript{35}, whereas counts per minute can differ considerably within LPA (from 101 up to 1500) and MVPA (≥1501).

The patient characteristics age, sex, weight, and height were recorded. The diagnosis was taken from the medical chart.

**Statistical analysis**

Intraclass correlation coefficients (ICC) were calculated for rest, LPA, MVPA, and PAL-based accelerometer and AD to analyze convergent validity. An ICC of ≥ 0.60 was rated as good validity, ≤ 0.3 to < 0.6 was rated as poor to moderate validity, and < 0.3 was rated as no convergent validity\textsuperscript{36}.

Differences between the AD and the accelerometer were analyzed using paired sample Student t-tests. Bland-Altman plots were drafted, where the difference between AD and accelerometer data was plotted against the average of both methods. Limits of agreement were calculated as mean difference ±1.96 * SD.
Differences between accelerometer and AD were analyzed using linear regression analysis for proportional bias\textsuperscript{37,38}.

The required number of measurement days to achieve an ICC of 0.75 and 0.9 for PAL, rest, LPA, and MVPA, measured with AD and accelerometer, were calculated using the Spearman-Brown prophecy formula: \( k = \frac{\text{ICC to achieve}}{1 - \text{ICC to achieve}} \times \frac{(1 - \text{ICC single})}{\text{ICC single}} \), where \( k \) is the number of required measurement days. Single-day ICC was calculated using repeated measurements of ANOVA by dividing the between-patient variance by the total variance, which is the sum of between-patient between-days and error variance. An ICC of > 0.75 was considered good reliability at group level, and an ICC of 0.9 was considered good reliability at an individual level\textsuperscript{39}.

Differences between rest, LPA, and MVPA measurements of the accelerometer with and without correction for non-wear were analyzed with paired sample Student t tests. Effect size was calculated by dividing the mean difference by the standard deviation of that difference. Bland-Altman plots were drawn, where the difference between rest, LPA, and MVPA based on accelerometer with and without correction for non-wear was plotted against the mean of both. Median difference was provided with limits of agreement as indicated by 2.5 and 97.5 percentiles. The Statistical Package for Social Sciences (IBM SPSS Statistics 22) was used for statistical analysis.

**RESULTS**

In 8 AD, parents and children described their activity precisely, but had not assigned a number to it. The authors filled in the number based on that description, consisting in total of 22 hours. Fourteen hours of missing values were imputed based on recall of parents and children. Thirteen diaries contained double numbers for a total of 17 hours. Missing values adjacent to smiley faces in the AD were imputed in 10 AD for a total of 8.5 hours. In 17 diaries, missing values were imputed by sitting activities for a total of 5 hours.

Eighty-three children participated in the Rheumates@Work study. After data imputation, 73 children (88%) had a complete AD. Sixty-six children (80%) had complete accelerometer data. In total, 61 children (73%) had a complete AD [consisting a
total of 10,248 hours, of which 21 hours were imputed (0.2%) and accelerometer data on 7 consecutive days (Table 1).

Table 1. Patient characteristics.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>10.1 (1.4)</td>
</tr>
<tr>
<td>Boys/girls, n</td>
<td>24/37</td>
</tr>
<tr>
<td>Height, cm</td>
<td>144 (10)</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>35.8 (8.7)</td>
</tr>
<tr>
<td>JIA subtype, n</td>
<td></td>
</tr>
<tr>
<td>Persistent oligoarticular</td>
<td>22</td>
</tr>
<tr>
<td>Extended oligoarticular</td>
<td>10</td>
</tr>
<tr>
<td>Polyarticular</td>
<td>17</td>
</tr>
<tr>
<td>Psoriasisrelated</td>
<td>3</td>
</tr>
<tr>
<td>Enthesitisrelated</td>
<td>3</td>
</tr>
<tr>
<td>Systemic</td>
<td>6</td>
</tr>
</tbody>
</table>

Values are mean (SD) unless otherwise specified.
JIA: juvenile idiopathic arthritis; cm: centimeters; kg: kilogram; n: number.

Number of hours spent on MVPA were mostly based on AD, followed by corrected accelerometer data and non-corrected accelerometer data (Table 2). Number of hours spent at rest was calculated mainly with accelerometer data (Table 2). The ICC between the AD and accelerometer indicate moderate convergent validity or worse (Table 2). The differences between AD and accelerometer depended on the means of both for all PA categories (Figure 1 and Table 3). For all PA categories for lower means, AD data were lower than those of the accelerometer, and in the higher means, AD data were higher than those of the accelerometer (Figure 1). Regression lines all run from lower left border to upper right border, indicating proportional bias.
Figure 1. Bland-Altman plots of differences in time spent in physical activity based on AD and accelerometer (Act) data (Y axis) plotted against the means of AD and accelerometer data (X axis). Solid horizontal lines: mean differences. Dotted lines: limits of agreement (LOA). (A) Minutes/day spent at rest (mean difference: -10, LOA -150.1; 130.3). (B) Minutes/day spent in LPA (mean difference: -16, LOA -173.9; 141.3). (C) Minutes/day spent in MVPA (mean difference: 26, LOA -50; 102.1). (D) PAL (mean difference: 0.09, LOA-0.15; 0.32). All regression lines run from lower left to upper right, indicating proportional bias. AD: activity diary; LPA: light physical activity; MVPA: moderate to vigorous physical activity; PAL: physical activity level.
Table 2. ICC between physical activity (minutes/day) and physical activity level based on data of AD and accelerometer, and differences between physical activity based on data of accelerometer and accelerometer corrected for non-wear.

<table>
<thead>
<tr>
<th>Variables</th>
<th>AD</th>
<th>Accelerometer</th>
<th>AD-Accelerometer, Mean Difference (95% CI), Significance*</th>
<th>ICC (95% CI)</th>
<th>Actical$_{cor}$</th>
<th>Actical$_{cor}$ - Accelerometer, Mean Difference (95% CI), Accelerometer, Significance*</th>
<th>Actical$_{cor}$ - Accelerometer, ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rest†</td>
<td>1156 (77)</td>
<td>1166 (53)</td>
<td>-10 (28.2 - 8.4), p=0.29</td>
<td>0.41 (0.19 - 0.60)</td>
<td>1145 (52)</td>
<td>-21(-25 to -16), p&lt;0.01</td>
<td>1.11</td>
</tr>
<tr>
<td>LPA†</td>
<td>209 (79)</td>
<td>225 (40)</td>
<td>-16 (-36.8 to 4.4), p=0.12</td>
<td>0.17 (-0.08 to 0.40)</td>
<td>243 (41)</td>
<td>18 (14 - 22), p&lt;0.01</td>
<td>1.12</td>
</tr>
<tr>
<td>MVPA†</td>
<td>76 (40)</td>
<td>50 (25)</td>
<td>26 (16.1 - 36.0), p&lt;0.01</td>
<td>0.24 (-0.01 to 0.46)</td>
<td>52 (25)</td>
<td>2 (0.9 - 3.8), p&lt;0.01</td>
<td>0.44</td>
</tr>
<tr>
<td>PAL</td>
<td>1.63 (0.14)</td>
<td>1.54 (0.09)</td>
<td>0.09 (0.05 - 0.11), p&lt;0.01</td>
<td>0.41 (0.09 - 0.63)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Values are mean (SD) unless otherwise specified. †Minutes/day. *Results of paired sample Student t test. ICC: intraclass correlation coefficient; AD: activity diary; LPA: light physical activity; MVPA: moderate to vigorous physical activity; PAL: physical activity level; Actical$_{COR}$: Actical (accelerometer) data corrected for non-wear time; ES: effect size (mean difference/ SD$_{difference}$); NA: not applicable.
Table 3. Results of the linear regression analyses to predict the difference between the AD and accelerometer, with the mean of the AD and accelerometer as predictors for assessing proportional bias.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Constant</th>
<th>b</th>
<th>95% CI</th>
<th>p</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rest</td>
<td>-578</td>
<td>0.49</td>
<td>0.18 - 0.80</td>
<td>&lt;0.01</td>
<td>0.15</td>
</tr>
<tr>
<td>LPA</td>
<td>-231</td>
<td>0.99</td>
<td>0.64 - 1.35</td>
<td>&lt;0.01</td>
<td>0.35</td>
</tr>
<tr>
<td>MVPA</td>
<td>-17</td>
<td>0.68</td>
<td>0.35 - 1.01</td>
<td>&lt;0.01</td>
<td>0.22</td>
</tr>
<tr>
<td>PAL</td>
<td>-0.83</td>
<td>0.57</td>
<td>0.31 - 0.84</td>
<td>&lt;0.01</td>
<td>0.25</td>
</tr>
</tbody>
</table>

AD: activity diary; LPA: light physical activity; MVPA: moderate to vigorous physical activity; PAL: physical activity level; CI: confidence interval.

To reach an acceptable reliability for determining MVPA, 5 days of accelerometer measurements were enough, and measured by means of AD, 13 days were necessary (Table 4). For clinical application for individual decision making, 14 days of measurement using the accelerometer and 36 days of measurement using the AD are needed (Table 4).

Table 4. ICC of AD and accelerometer, and number of days to reach an ICC of 0.75 and 0.90.

<table>
<thead>
<tr>
<th>Activity</th>
<th>ICC (95% CI)</th>
<th>No. days needed to reach:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ICC 0.75</td>
</tr>
<tr>
<td>Activity diary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAL</td>
<td>0.21 (0.12 - 0.32)</td>
<td>11.4</td>
</tr>
<tr>
<td>Rest</td>
<td>0.32 (0.23 - 0.44)</td>
<td>6.3</td>
</tr>
<tr>
<td>LPA</td>
<td>0.36 (0.26 - 0.48)</td>
<td>5.3</td>
</tr>
<tr>
<td>MVPA</td>
<td>0.20 (0.11 - 0.31)</td>
<td>12.3</td>
</tr>
<tr>
<td>Accelerometer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAL</td>
<td>0.33 (0.23 - 0.47)</td>
<td>6.2</td>
</tr>
<tr>
<td>Rest</td>
<td>0.37 (0.27 - 0.49)</td>
<td>5.1</td>
</tr>
<tr>
<td>LPA</td>
<td>0.37 (0.27 - 0.49)</td>
<td>5.2</td>
</tr>
<tr>
<td>MVPA</td>
<td>0.39 (0.29 - 0.51)</td>
<td>4.6</td>
</tr>
</tbody>
</table>

ICC: intraclass correlation coefficient; PAL: physical activity level; LPA: light physical activity; MVPA: moderate to vigorous physical activity; CI: confidence interval.

When accelerometer data were corrected for non-wear, significant differences were found in mean time spent at rest, LPA, and MVPA (Table 2). The effect size for MVPA was small for non-wear. For rest and LPA, effect sizes were large.
Bland-Altman plots showed that corrections for non-wear of the accelerometer in MVPA could differ up to 25 minutes for individual patients (Figure 2).
DISCUSSION

Our study showed that the AD and accelerometer have a moderate or poor convergent validity in patients with JIA aged 8 to 13 years. One-week measurement with an accelerometer and 13 days of measurements with an AD are sufficient to obtain reliable estimates of PA at group level. In individual cases and for clinical purposes, almost 3 weeks of accelerometer and > 5 weeks of AD measurements are required. Correction for non-wear of the accelerometer resulted in a significant increase in LPA and MVPA. The effect size for LPA was substantial, and for MVPA it was small. In studies where LPA is one of the outcome variables, correcting for non-wear is relevant. Correcting MVPA for non-wear is relevant for individual patients.

The poor to moderate convergent validity between AD and accelerometer was also found previously. Two-thirds of parents of healthy children, aged between...
5 to 7 years, overestimated their children’s PA when PA was measured with an electronic diary compared to the activity counts of the accelerometer. A moderate correlation (0.44) was found between both instruments\textsuperscript{41}. Correlations of 0.33 for girls and 0.44 for boys were found, controlled for body mass, between estimated AEE measured with a 3-day AD and an accelerometer in 403 healthy adolescents\textsuperscript{40}. In healthy Spanish adolescents, a moderate correlation of 0.36 was found in MVPA between the ActiGraph brand activity monitor and the Bouchard AD\textsuperscript{42}. In general, correlations between any self-report and an objective instrument were found to be low to moderate at best\textsuperscript{25}.

The poor to moderate convergent validity can be explained in 2 ways. First, in AD, participants usually tend to overestimate the intensity and duration of different types of activities and sports because of the intermittent characteristics of activities and sports\textsuperscript{25}. When a child reports 1 hour of physical education classes, normally classified as MVPA, in reality only 37% of the time will be actual MVPA while the rest of the time will be spent on sedentary or LPA\textsuperscript{26}. Additionally, accelerometers underestimate intensity and duration of certain types of activities because they are less sensitive to registering activities such as walking up stairs, cycling, and activities that mainly involve arm movements\textsuperscript{43}. Moreover, compliance with wearing an accelerometer for a whole 7-day period remains a concern, and non-wear will again underestimate PA\textsuperscript{43}. Second, intensity thresholds of AD are based on metabolic equivalents of tasks performed, while thresholds of accelerometers are measured in the laboratory, where body movement and energy expenditure are concurrently measured\textsuperscript{25}.

For children with JIA, disease-specific arguments may account for the poor to moderate convergent validity between AD and accelerometer. Children with JIA have higher AEE compared with healthy peers when performing similar activities\textsuperscript{44,45}. This difference not only affects thresholds for activity counts for categories of PA, but it also affects the classification of the activities 1-9, as used in the AD. Additionally, children with JIA have different activity counts compared with healthy peers when performing the same activities\textsuperscript{44}. To what extent both arguments affect convergent validity has not been studied, to our knowledge, but should be taken into account.
Our study showed that in children with JIA aged 8-13 years, 1 week of measurements with an accelerometer was sufficient, but for an AD, at least 13 days of measurements were needed to reach sufficient reliability. This finding is in line with a previous study that showed that the number of measurement days for reliable assessment of PA depended on the type of instrument, purpose of the study, and the characteristics (including age) of the population. Healthy younger children exhibited less day-to-day variability than healthy adolescents and therefore required fewer days to assess PA reliably. In healthy 5-year-old preschool children, 5-6 days of accelerometer monitoring were needed, compared to 4-5 days in 7-to 12-year-old children, and 8-9 days of monitoring in 13- to 16-year-olds. For adults, 3-5 days of monitoring appeared to be sufficient to assess PA. In healthy and chronically ill children, as far as we know, the number of days required for the AD has never been assessed. Our results indicate that in children with JIA on an individual level and for clinical purposes, almost 3 weeks of accelerometer monitoring is needed and 5 weeks of the AD. This number of weeks is not realistic, considering the effort this would require from children and their parents.

We found a significant but small increase of about 4% between MVPA, measured with and without correction for non-wear. In a study including 513 healthy children, aged 13-15 years, correction for non-wear using ActiGraph accelerometers and a non-wear diary resulted in an increased mean MVPA of 43% (23-33 minutes/day increase). In the study, the increase was mainly related to non-wear during aquatic activities and ball games. In our study, a smaller correction for non-wear for MVPA was found, perhaps because of the instructions given to the children to wear the accelerometer all the time, except for water activities. Another explanation might be the younger age of our patients, which could lead to more compliance to wear the accelerometer. However, at individual level, MVPA was corrected for 10 minutes up to 25 minutes per day in 5 children, and for 5–10 minutes in 6 patients. This correction is clinically relevant because it results in an increase of 35-175 min of MVPA per week. These findings indicate that the relevance of correcting for non-wear can vary between samples and that, in studying MVPA in JIA clinical trials, correction leads to small differences at the group level. For clinical use in children with JIA individually, the use of an AD in combination with an accelerometer is recommended, because in individual cases non-wear can be considerable.
Our study has limitations. Only children with JIA with no or mild disease activity were selected. Patients with high disease activity may show lower and less variable PA, and are more likely to engage in exercise activities such as swimming, resulting in more non-wear of the accelerometer. However, measuring PA in low disease activity states is more useful, because it is during this phase that PA is especially resumed. Another form of selection bias was caused by the willingness of children to participate in a program aimed at improving PA. These children may have overestimated their PA, leading to higher AD scores, or those who were less active were willing to improve their PA level. In our study, boys were relatively more represented as compared with the general population of patients with JIA. Boys may have different activity patterns that could have influenced our results. The age of the patients may have also influenced results. Children in our study were 8-13 years old, but the reliability of the AD has only been assessed in children 10 years of age and older\textsuperscript{32}. We tried to overcome this by instructing parents to help their children fill in the diary. Another limitation was that the AD was validated only in children aged 15 years\textsuperscript{33}. Imputing missing values could cause errors, although this was only necessary in a very small proportion of the AD. An epoch of 1 minute for the accelerometer could be another limitation because it underestimated MVPA in preschool children and adolescents compared with an epoch of 15 seconds\textsuperscript{48,49}. In a recent study in healthy children, aged 8-11 years, a small clinically irrelevant underestimation of MVPA (1.9 minutes/day) was found when using an epoch of 1 minute\textsuperscript{50}. Another limitation is that we measured for 7 days and used these data to calculate the number of days needed for reliable estimates. By measuring and using a single ICC, compound symmetry is assumed, meaning that the correlations among days are similar\textsuperscript{24}. However, because of day-to-day variability, actual correlations between days will most likely differ, thus violating the compound symmetry assumption leading to underestimation of the days required\textsuperscript{24}.

There is poor to moderate convergent validity between the AD and accelerometer. To compare PA between groups of patients with JIA, a 1-week assessment with the accelerometer is sufficient. For individual decision making, 2-3 weeks are required. To be able to correct for non-wear (for instance swimming), use of an AD is recommended.
REFERENCES

Chapter 5

APPENDIX

Appendix 1. Categories of activities and the formulas for energy expenditure.

Categories of activities for the activity diary
1= sleeping or resting in bed; 2= sitting, eating, writing, etc.; 3= standing, washing, combing, etc.; 4= walking indoors (< 4 km/hour), light home activities; 5= walking outdoors (4-6 km/hour), cleaning bedroom, easy outdoor playing; 6= recreational sports and leisure time activities with low intensity; 7= recreational sports and leisure time activities with moderate intensity; 8= recreational sports and leisure time activities with high intensity; and 9= sports competitions.

Equations to calculate energy expenditure
Rest time refers to activities that do not increase energy expenditure substantially above resting level, such as sleeping, lying down, and seated activities. These are represented by categories 1 and 2, and the energy costs are 0.98*basal metabolic rate (BMR) and 1.5*BMR, respectively. Intensity thresholds between light physical activity (LPA) and moderate to vigorous physical activity (MVPA) are around 4 metabolic equivalents of tasks. Therefore, LPA is represented by categories 3, 4, and 5, with a cost of 2.0, 2.8, and 3.3*BMR, respectively. MVPA is category 6 and higher, with an energy cost of 4.4, 6.5, 10.0, and 15.0*BMR, respectively.

ACKNOWLEDGEMENT
We thank Marion van Rossum, Jeannette Capponn, Marco van Brussel, and Nico Wulffraat for selecting the patients and for their assistance in the data collection. We also thank the language and translation center of the University of Groningen for language editing.
Measuring physical activity in JIA