

Article

Machine Learning-Based Aggression Detection in Children with ADHD Using Sensor-Based Physical Activity Monitoring

Catherine Park ¹, Mohammad Dehghan Rouzi ¹ , Md Moin Uddin Atique ¹, M. G. Finco ¹ , Ram Kinker Mishra ¹, Griselda Barba-Villalobos ², Emily Crossman ², Chima Amushie ², Jacqueline Nguyen ², Chadi Calarge ^{2,*} and Bijan Najafi ^{1,*} 

- ¹ Interdisciplinary Consortium on Advanced Motion Performance (iCAMP), Michael E. DeBakey Department of Surgery, Baylor College of Medicine, Houston, TX 77030, USA; catherine.park@bcm.edu (C.P.); mohammad.dehghanrouzi@bcm.edu (M.D.R.); moindu44@gmail.com (M.M.U.A.); graci.finco@bcm.edu (M.G.F.); ram.mishra@biosensics.com (R.K.M.)
- ² Menninger Department of Psychiatry and Behavioral Sciences and Department of Pediatrics, Baylor College of Medicine, Texas Children's Hospital, Houston, TX 77030, USA; griselda.barbavillalobos@bcm.edu (G.B.-V.); emily.crossman@bcm.edu (E.C.); chima.amushie@bcm.edu (C.A.); jacqueline.nguyen@bcm.edu (J.N.)
- * Correspondence: chadi.calarge@bcm.edu (C.C.); bijan.najafi@bcm.edu (B.N.); Tel.: +1-832-824-4764 (C.C.); +1-713-798-7536 (B.N.)

Abstract: Aggression in children is highly prevalent and can have devastating consequences, yet there is currently no objective method to track its frequency in daily life. This study aims to investigate the use of wearable-sensor-derived physical activity data and machine learning to objectively identify physical-aggressive incidents in children. Participants ($n = 39$) aged 7 to 16 years, with and without ADHD, wore a waist-worn activity monitor (ActiGraph, GT3X+) for up to one week, three times over 12 months, while demographic, anthropometric, and clinical data were collected. Machine learning techniques, specifically random forest, were used to analyze patterns that identify physical-aggressive incident with 1-min time resolution. A total of 119 aggression episodes, lasting 7.3 ± 13.1 min for a total of 872 1-min epochs including 132 physical aggression epochs, were collected. The model achieved high precision (80.2%), accuracy (82.0%), recall (85.0%), F1 score (82.4%), and area under the curve (89.3%) to distinguish physical aggression epochs. The sensor-derived feature of vector magnitude (faster triaxial acceleration) was the second contributing feature in the model, and significantly distinguished aggression and non-aggression epochs. If validated in larger samples, this model could provide a practical and efficient solution for remotely detecting and managing aggressive incidents in children.

Keywords: pediatrics; aggression; wearables; remote patient monitoring; machine learning



Citation: Park, C.; Rouzi, M.D.; Atique, M.M.U.; Finco, M.G.; Mishra, R.K.; Barba-Villalobos, G.; Crossman, E.; Amushie, C.; Nguyen, J.; Calarge, C.; et al. Machine Learning-Based Aggression Detection in Children with ADHD Using Sensor-Based Physical Activity Monitoring. *Sensors* **2023**, *23*, 4949. <https://doi.org/10.3390/s23104949>

Academic Editor: Ki H. Chon

Received: 12 April 2023

Revised: 10 May 2023

Accepted: 20 May 2023

Published: 21 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Aggressive incidents, which range from verbal expressions of anger to physical violence [1], are the most common reason for referral to child and adolescent psychiatric services and the primary indication for inpatient psychiatric admission [2–7]. Childhood aggression is often comorbid with other psychiatric disorders, including attention deficit hyperactivity disorder (ADHD), disruptive behavior disorders, mood disorders, anxiety disorders, psychotic disorders, trauma-based disorders, tic disorders, intellectual disability, and autism spectrum disorder [2–6,8]. Aggressive incidents in children have been associated with loss of life, as well as billions of dollars in decreased productivity, quality of life, and healthcare costs [9–11].

Despite the high prevalence and devastating consequences of aggressive incidents in children with ADHD, to our knowledge, no objective method exists to track their frequency and severity in daily life. Available tools examine dimensions of aggression (e.g., impulsive vs. premeditated motivations, primary targets, severity, and duration),

but they are subjective, typically relying on the parent or caregiver [12–15]. Consequently, intervention studies targeting aggressive incidents have been largely restricted to rating scales to capture treatment response [16–19]. While the face validity of the available rating scales is high, accuracy may be suboptimal because informants may have limited reliability. For instance, a parent/caregiver may not be fully informed about the child's behavior, given the substantial time the child spends at school. A valid and objective method to capture aggressive incidents in daily life could help inform care for children with aggression to ultimately optimize treatment, improve quality of life, and reduce associated costs. Notably, wearable sensors or smart technologies (e.g., smart watches) could offer practical, cost-effective solutions to objectively measure aggressive incidents in children.

Several studies have demonstrated the efficacy of using wearable sensors to measure physical activity in children with psychopathology. For instance, accelerometers and inertial measurement units have been used to characterize physical activity, sleep patterns, and motor coordination in children with ADHD and other neurological disorders [20–24]. However, these studies focused on task-oriented activities, rather than activities of daily living. Two recent studies used a smartwatch to analyze the movements of children with ADHD at school [25] and while performing daily activities (e.g., sitting, exercises, and household activities) [26]. Notably, Lindhiem et al. used a machine learning approach to establish the context in which hyperactivity is present [26]. As for using wearables specifically to quantify aggressive incidents, Goodwin et al. found that physiological and motion data collected with a biosensor worn by 6 to 17 year old children with autism spectrum disorder admitted to an inpatient psychiatric unit allowed the identification of aggression during observation sessions [27]. However, to our knowledge, this work has not been extended to a naturalistic outpatient setting.

This study aims to address the existing gap by devising an innovative method for identifying aggressive incidents in children through the use of wearable technology and machine learning. Our objectives are to: (1) assess the efficacy of machine learning in detecting aggressive incidents, (2) uncover distinct physical activity patterns during aggressive incidents, and (3) pinpoint the optimal sensor-driven features, also known as “digital biomarkers,” for identifying aggressive incidents. By utilizing physical activity data gathered from wearable sensors, we hypothesize that our approach will provide a valid and dependable means of detecting aggressive incidents in children.

2. Materials and Methods

2.1. Participants

Data for this analysis were collected in the context of a year-long, observational study examining the skeletal effects of psychostimulants. Participants were recruited from pediatric outpatient clinics. Eligible participants were medically healthy males and females who were 7 to 16 years old, of all racial and ethnic backgrounds, within one month of starting a psychostimulant or unmedicated, with no plans to move out of state for one year. Potential participants were excluded if they: (1) had treatment with psychotropics or other medications affecting bone metabolism within the prior year, (2) had serious medical conditions involving a vital organ or affecting bone metabolism, or pregnancy/lactation, (3) were underweight or overweight (i.e., body mass index less than the 5th percentile or greater than the 95th percentile), (4) had a substance use disorder or an eating disorder, or (5) had an intellectual disability or an inability to understand the English language, which would impair their ability to complete study procedures. The local Institutional Review Board approved the study, and written informed consent and verbal assent were obtained.

After completing the baseline visit, participants returned for an in-person visit at 6 and 12 months and completed a monthly phone visit in between. At every contact, the medical and treatment history was updated, including psychostimulant use. In addition, medication adherence was assessed. At in-person visits, demographic and anthropometric variables were collected, including age, sex, height, and weight. The parent also completed the child behavior checklist (CBCL), a widely used rating scale with excellent validity

and reliability [28,29]. The CBCL generates a dimensional profile composed of several problem scales, including rule-breaking and aggressive behavior scales. Extensive norms are available, allowing for generation of T-scores. The study psychiatrist also completed the clinical global impression scale (CGI-Severity), a single-item scale with scores ranging between 0 (not assessed) and 7 (among the most extremely ill patients), with a score ≥ 3 typically considered clinically significant [30]. Clinical diagnoses were based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [31] and incorporated information from the medical record, the rating scales, and an unstructured interview of the parent/guardian and the child completed by a child and adolescent psychiatrist.

Height and weight were used to calculate body mass index (BMI, kg/m^2), and sex-age-specific BMI Z-scores were derived [32]. Where applicable, the daily psychostimulant dose was converted to a methylphenidate (MPH) equivalent [33] and the daily average weight- and adherence-adjusted dose in MPH equivalent was computed. The T-score of the CBCL factors was dichotomized based on a threshold of ≥ 60 to identify participants at risk for clinically significant behavioral problems [29].

Because parents may not be aware of their child's aggressive incidents on days when school is in session, days when physical activity was collected were classified as "school" vs. "non-school" days. The latter included weekends and holidays.

2.2. Assessment of Physical Activity and Aggression Episodes

At the in-person visits, participants were provided with a waist-worn accelerometer sensor (GT3X+, ActiGraph Corp., Pensacola, FL, USA), which they were instructed to wear for seven consecutive days starting from the day following their visit (Figure 1). The ActiGraph GT3X+ is compact (4.6 cm (W) \times 3.3 cm (H) \times 1.5 cm (D)), lightweight (19 g), user-friendly, and has a battery life of up to 25 days. In addition, the parents were given a log, based on the Retrospective Modified Overt Aggression Scale (R-MOAS) [34], and asked to record (during the same seven-day period) the date, time, and duration of both verbal-aggressive incidents and physical-aggressive incidents. Given that physical aggression incidents are more problematic and can be reported more validly, this study concentrated on the physical-aggressive incidents, which included incidents towards other people, incidents involving property, and incidents directed at self.

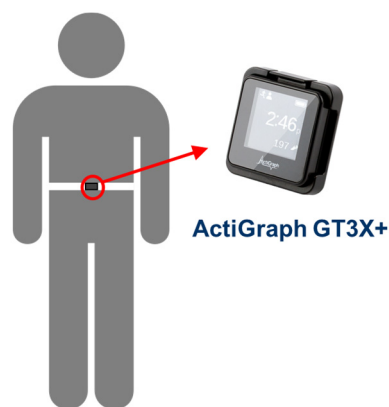


Figure 1. A wearable sensor (GT3X+, ActiGraph Corp., Pensacola, FL, USA) and its placement for physical activity assessments.

Information from the R-MOAS allowed definition of the aggression periods, each comprising one or more continuous aggressive incidents.

2.3. Sensor-Driven Features

To quantify physical activity, seven variables (vector magnitude, cadence, % standing, % sitting, % lying, kilocalories, and MET rate; Table 1) were calculated, using ActiLife software based on 1-min-long epochs (ActiGraph, Pensacola, FL, USA) [35,36]. Vector

magnitude indicates the time rate of change of movement speed [37]. Kilocalories indicates the amount of energy relative to body mass and vector magnitude [38]. Metabolic rate (i.e., MET rate) indicates the rate of energy expended per unit of time.

Table 1. Sensor-derived physical activity variables, measured in 1-min epochs. The physical activity variables are obtained from a waist-worn accelerometer and reflect the intensity and level of physical activity in 1-min intervals. These variables are utilized to identify and describe aggression epochs, and they are also one of the inputs to the machine learning model.

Name	Unit	Description
Vector magnitude (vm)	cpm	Total vector magnitude of three accelerometer axes, calculated by the ActiGraph using the following equation [37]: $vm = \sqrt{axis1^2 + axis2^2 + axis3^2}$
Cadence	steps/min	Step counts per minute
% Standing	% of epoch	Percentage of an epoch when participant was standing
% Sitting	% of epoch	Percentage of an epoch when participant was sitting
% Lying	% of epoch	Percentage of an epoch when participant was lying
Kilocalories	cpm	Kilocalories expended per minute, calculated by the Actigraph using the following equation [38]: $\frac{kcal}{min} = (0.001064 \times vm) + (0.087512 \times \text{body mass in kg}) - 5.500229$
MET Rate	cpm	Metabolic equivalents to measure energy expenditure

cpm: counts per minute. MET: metabolic rate.

2.4. Random Forest Classifier Approach for Feature Selection

Random forest uses ensemble learning, combining multiple classifiers to provide solutions to complex problems [39,40], and is also commonly used for feature importance determination [39]. As such, our approach used a random forest algorithm both for feature selection and the construction of a classification model, seeking to rank demographic, clinical, sensor-derived physical activity, and observed day type variables with regard to their potential to capture aggressive incidents. To do so, each 1-min epoch was categorized as aggression or non-aggression, based on whether it occurred during a reported time period of aggression (based on the R-MOAS) or not, respectively. All aggression epochs were included in the model. As expected, the number of non-aggression epochs was much larger, causing unbalance in the dataset. To address this issue [41], an equal number of non-aggression epochs was included, selected randomly from all available non-aggression epochs with the following restrictions: (1) non-aggression epochs derived from sleep time (i.e., 10 p.m.–6 a.m.) and (2) those occurring within 30 min before and after an aggression epoch, to minimize any potential contamination effect (Figure 2).

The random forest classifier algorithm included several features, including age, sex, sex-age-specific BMI and height Z-scores, the weight-adherence-adjusted dosage of psychostimulant (in MPH equivalent, mg/kg), whether the epoch occurred on a school day, illness severity based on the CGI, and T-scores for six CBCL factors: total, internalizing, externalizing, attention problems, rule breaking, and aggressive behavior. Additionally, the algorithm considered seven sensor-derived variables (see Table 1). By including both demographic and clinical factors, as well as sensor-derived variables, the algorithm can provide a more nuanced understanding of the relationship between these factors and aggressive incidents. The sensor-derived variables, in particular, are important as they offer an objective measure of physical activity and movement patterns, potentially related to the occurrence of aggressive incidents.

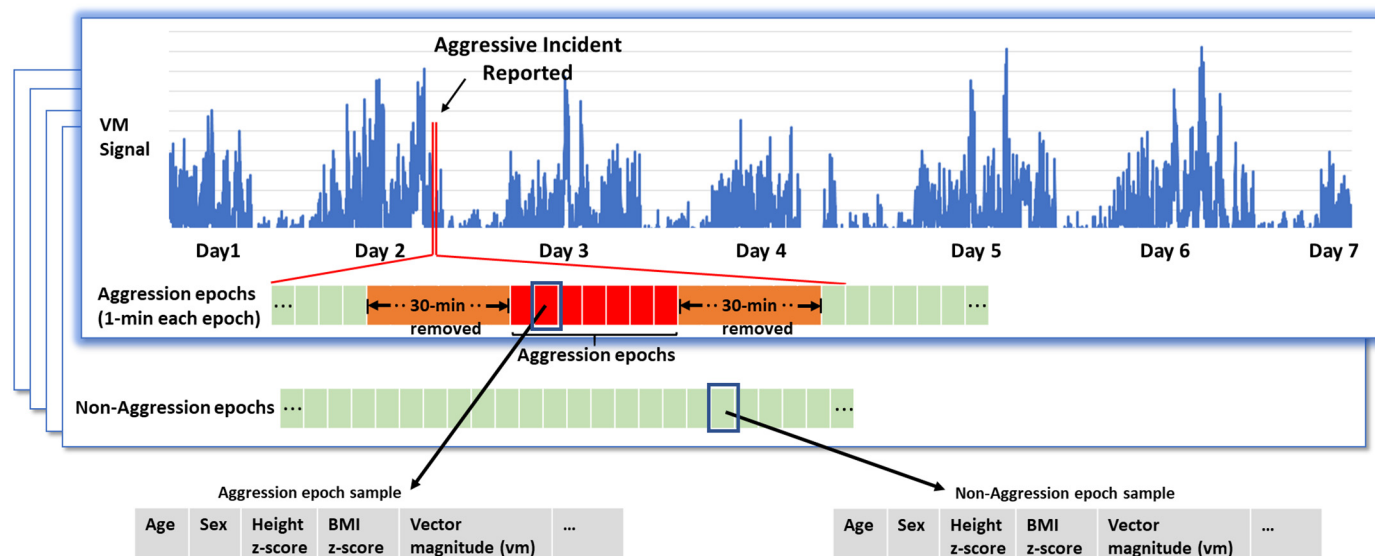


Figure 2. (Top): An example of the vector magnitude (vm) signal, which is one of the sensor-derived physical activity variables, for both aggression and non-aggression epochs. Each rectangle represents an epoch, consisting of 1-min worth of physical activity data. The red-filled boxes indicate that the epoch occurred during a reported aggression episode, while the green ones indicate the absence of aggression. The large orange rectangles represent 30-min worth of data occurring before and after an aggressive incident. These were excluded from the analyses to minimize contamination risk. **(Bottom):** This diagram illustrates the different types of variables that are associated with each epoch for inclusion in the model. These variables include demographic, clinical, sensor-derived physical activity, and observed day type variables.

Figure 3 depicts a flowchart divided into two steps. The initial step involved data cleaning, analysis, and preparation for the machine learning model, while the second step detailed the model's design. In the primary step, sensor data were labeled based on parent-reported physically aggressive incidents, and clinical and demographic information were integrated into each epoch. Subsequently, epochs collected during sleep and within 30 min before and after aggressive incidents were excluded. The second step comprised a two-step feature selection process. The first step was to evaluate the importance of all sensor-derived features using a random forest classifier. We used training and testing sample sizes of 70% and 30%, respectively. The random forest model was trained with 500 trees with a balanced subsample for class weights to improve performance [42,43]. We also used a grid search cross-validation technique to optimize three attributes of the random forest model, an effective approach for optimizing the performance of classifiers [44,45].

In the second step, we ran a random forest method with the same training and testing sizes of 70% and 30%, respectively. This procedure was repeated for 100 bootstraps to ensure that all observations were selected in the validation sub-sample, allowing computation of the means and their standard deviations to quantify their uncertainties [46]. Each step involved adding one feature to the dataset and training the model to gain insight into how adding each feature affects the model's performance. The final reports comprised the mean and standard deviation (SD) of all 100 estimated model performance metrics, including the accuracy, recall, precision, F1 score, and area under the curve (AUC). Feature selection processes were conducted using Python version 3.10 (Python Software Company, Fredericksburg, VA, USA).

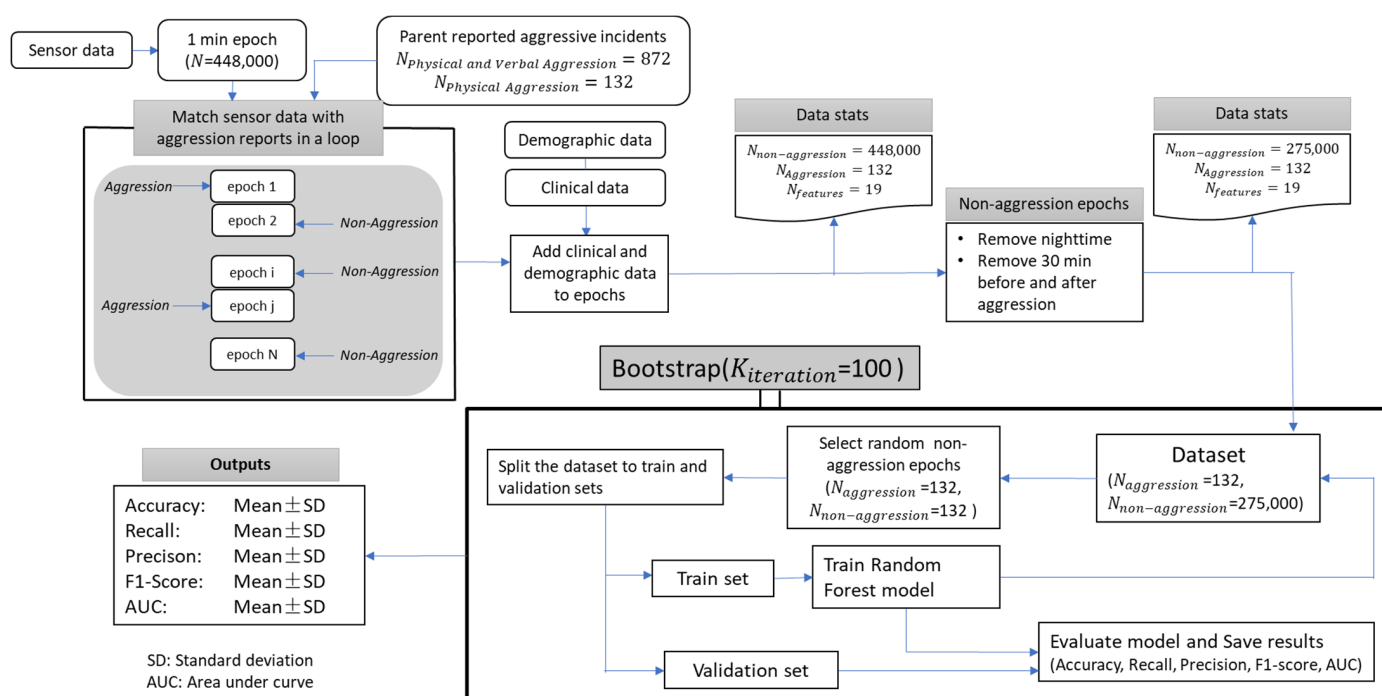


Figure 3. This flowchart illustrates the data preprocess and feature selection using a machine learning model, as well as the evaluation of the resulting model. The random forest classifier model was employed, with 100 bootstraps utilized to assess model performance. The evaluation metrics included mean \pm SD values for model accuracy, recall, precision, F1 score, and area under the curve (AUC).

2.5. Statistical Analysis

Continuous data are expressed as mean \pm standard deviation (SD), while categorical data are reported as count (%). The Shapiro–Wilk test was used to determine whether continuous variables followed a normal distribution. The one-way ANOVA for normally distributed variables and the Mann–Whitney U test for non-normally distributed variables were used to compare mean differences between epochs with vs. those without aggression. The chi-square test was used to examine between-group differences in categorical variables. All statistical analyses were conducted using IBM SPSS Statistics version 27 (IBM Corp., Armonk, NY, USA).

3. Results

3.1. Demographic and Clinical Characteristics

Table 2 presents the demographic and clinical characteristics of participants who had at least one time period of aggression. There were a total of 31 time periods of physical aggression, each lasting an average of 4.3 ± 7.8 min, resulting in 132 physical aggression epochs. Figure 4A depicts the analysis of the duration of reported aggression (verbal and physical) as a function of time of day, including night time (10 p.m.–6 a.m.), morning time (6 a.m.–12 p.m.), and afternoon/evening (12 p.m.–10 p.m.) (time frames selected for purposes of visualization). Time periods of physical aggression mostly occurred after school and were most pronounced around 6 p.m. Time periods of verbal aggression had two pronounced peaks at 6 a.m. and 3 p.m. Furthermore, the majority of physical aggression occurred during weekend days, particularly on Sundays (Figure 4B).

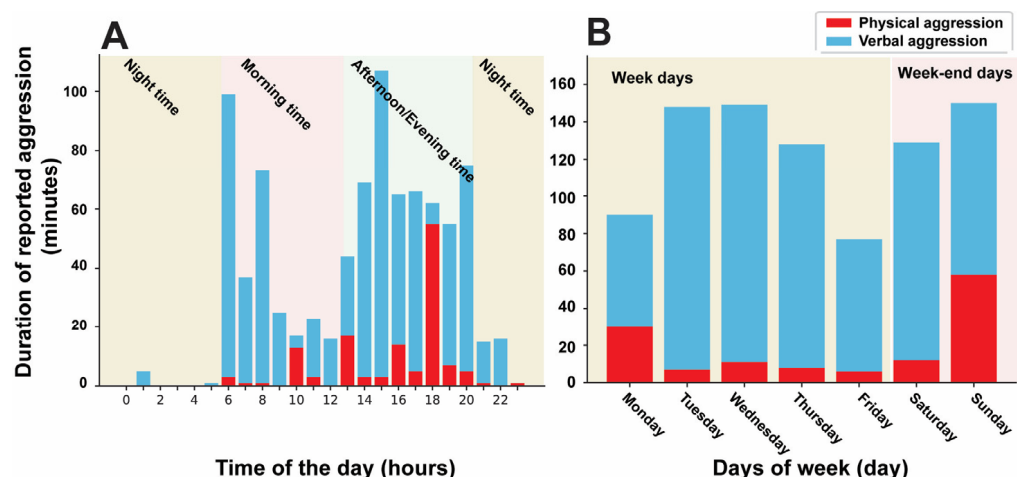


Figure 4. (A) The duration of reported aggression (verbal and physical) as a function of time of day. (B) The duration of reported aggression (verbal and physical) as a function of weekdays and weekend days. The majority of physical aggression epochs occurred during the afternoon and evening, and on weekend days.

Table 2. Demographic and clinical characteristics.

Variables	N = 39
Demographics	
Age, years	9.4 ± 1.7
Sex (Male), n (%)	31 (79.5%)
Hispanic, n (%)	13 (33.3%)
Race, (%)	
White	29 (74.4%)
Black	5 (12.8%)
Native American	1 (2.6%)
Others	4 (10.3%)
Height (cm)	136.2 ± 12.6
Height Z-score	0.008 ± 0.893
Weight (kg)	32.1 ± 9.2
Body Mass Index, kg/m ²	17.0 ± 1.9
Body Mass Index Z-score	0.1 ± 0.8
Clinical Characteristics	
Medication, N (%)	34 (87.2%)
Daily Average Weight and Adherence Adjusted	0.51 ± 0.37
Dose in MPH Equivalency (MPH Eq mg/kg)	
CBCL, Total Score	44.2 ± 10.4
Aggressive Behavior	60.7 ± 9.1
Rule-Breaking Behavior	57.4 ± 7.6
Attention Problems	64.1 ± 9.1
Internalizing Problems	54.1 ± 12.1
Externalizing Problems	58.1 ± 10.9
CGI-Severity Score	4.7 ± 1.0

Values are presented as mean ± standard deviation or n (%). All CBCL components are reported in T-Scores. MPH: methylphenidate; CBCL: child behavior checklist; CGI-Severity: clinical global impression-severity.

3.2. Difference in Features between Aggression and Non-Aggression Epochs

Table 3 presents descriptive statistics for the physical activity and observed day type features of both aggression and non-aggression epochs collected from 39 participants with ADHD (n = 35) and without (n = 4). Movement acceleration (accelerometer vector magnitude) was significantly higher during the former.

Table 3. Comparison of sensor-derived body motion and posture characteristics and observed day type between physical aggression and non-aggression epochs. The term “non-aggression epochs” represents periods where no physical aggressive incident was observed, while “aggression epochs” refers to the time periods during which a physical-aggressive incident was recorded. The table displays the mean values and standard deviations of each variable for both aggression and non-aggression epochs, as well as the *p*-value from the statistical test of the difference between the two groups.

Variables	Aggression Epochs N = 132	Non-Aggression Epochs N = 132	<i>p</i> -Value
Physical activity			
Vector magnitude (vm)	1580.7 ± 1831.1	873.3 ± 1137.2	0.027 *
Cadence (cpm)	15.3 ± 22.5	8.1 ± 13.6	0.102
% Standing	51.4 ± 47.7	44.6 ± 46.2	0.320
% Sitting	16.9 ± 34.2	26.4 ± 41.9	0.176
% Lying	15.7 ± 36.0	14.9 ± 35.7	0.478
Kilocalories (cpm)	0.27 ± 0.46	0.17 ± 0.23	0.427
MET rate (cpm)	1.15 ± 0.43	1.04 ± 0.12	0.132
Observed day type			
Non-school day	47.0%	47.7%	0.902

Values are presented as mean ± standard deviation or n (%). Asterisk denotes a significant difference between the groups. cpm: counts per minute. MET: metabolic rate.

3.3. Optimal Feature Selection and Evaluation

Figure 5A displays the ranking of the 20 features examined, using the random forest classifier algorithm. A higher percentage value indicates that the feature had a greater contribution to the model. The top 10 features that contributed most to the model, in order of importance, were age (14.41%), vector magnitude (9.27%), CBCL Total T-score (7.96%), height Z-score (7.16%), adjusted MPH dose (6.57%), BMI Z-score (5.56%), CBCL attention T-score (5.37%), CBCL aggressive T-score (5.12%), CBCL internalizing T-score (4.97%), and CBCL externalizing T-score (4.70%).

Figure 5B shows the model validation results as AUC, F1, accuracy, recall, and precision as a function of the number of ranked features. The model with all 20 features achieved an AUC of $89.3 \pm 4.2\%$, F1 of $82.4 \pm 4.7\%$, accuracy of $82.0 \pm 4.8\%$, recall of $85.0 \pm 6.6\%$, and precision of $80.2 \pm 5.3\%$. For the excellent range of AUC and F1 (0.8 to 0.9) and acceptable range of accuracy, recall, and precision (0.7 to 0.8), age, vector magnitude, CBCL Total T-score, and height Z-score were identified as the most important features. A model using these four features achieved an AUC of $87.2 \pm 4.1\%$, F1 score of $80.8 \pm 4.3\%$, accuracy of $80.1 \pm 4.6\%$, recall of $83.8 \pm 6.7\%$, and precision of $78.2 \pm 5.9\%$. When the CBCL Total T-score feature was excluded from the model, the performance decreased by about 10%, resulting in an AUC of 77.9 ± 5.2 , an F1 score of 71.8 ± 5.0 , an accuracy of 71.6 ± 4.8 , a recall of 71.7 ± 8.3 , and a precision of 71.9 ± 5.4 . However, these values still fall within an acceptable range. It should be noted that the collection of CBCL values is often resource-intensive, and their exclusion may make the model more practical for implementation in certain settings.

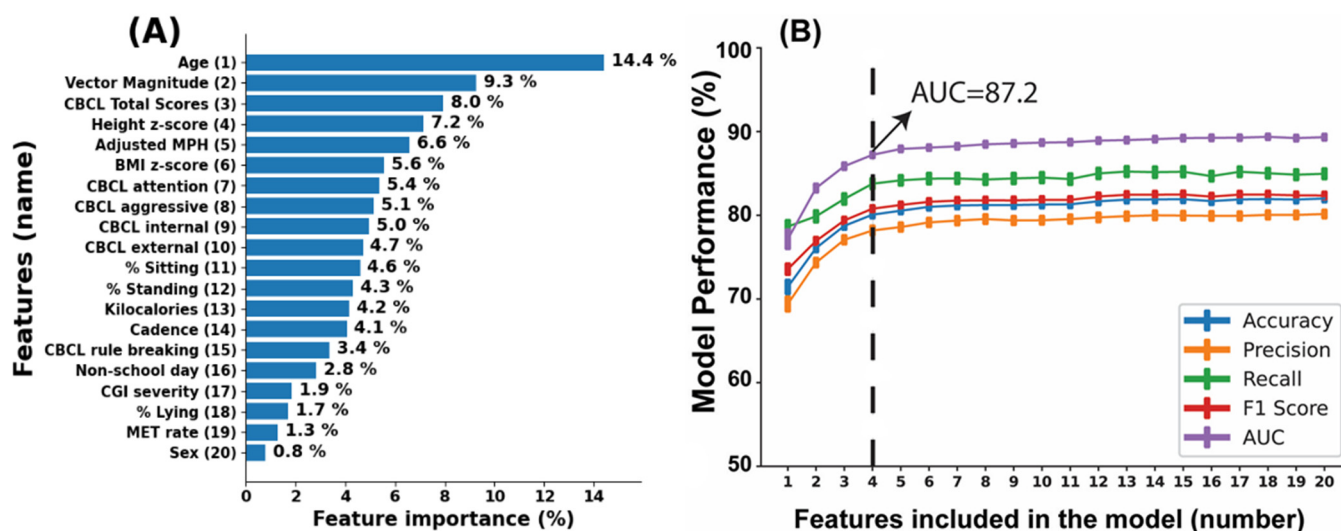


Figure 5. Figure (A) shows the ranking of 20 features based on their significance in distinguishing physical aggression from non-aggression epochs, as determined by the random forest classifier algorithm. Meanwhile, Figure (B) demonstrates the model's effectiveness in distinguishing between the two groups, as measured by AUC, F1 score, accuracy, recall, and precision. The figure utilizes several abbreviations, such as CBCL (child behavior checklist), MPH (methylphenidate), CGI-S (clinical global impression-severity), AUC (area under curve).

4. Discussion

This study aimed to investigate whether sensor-derived physical activity data would optimize the identification of aggressive incidents in children, using machine learning. First, this approach achieved a high performance. Second, four features disproportionately accounted for the excellent model performance. Third, three of these top features were objective measures, including age, height, and sensor-based vector magnitude (triaxial acceleration). Finally, excluding the one measure based on parent report (i.e., the CBCL Total Score) reduced model performance, though not drastically.

Identifying aggression epochs is a crucial step in preventing and managing aggressive incidents. However, unlike the vast majority of the work undertaken so far, which has relied on subjective report, our study suggests that objective assessment is feasible, with potentially high reliability. Importantly, a wearable technology-based measure, specifically vector magnitude, appears to comprise one of the top four features contributing to the model's excellent performance. This feature has face validity, given that physical aggression often involves rapid triaxial acceleration in order to commit the physically violent act. It, then, follows that tracking vector magnitude could be useful in identifying the onset of aggressive incidents. Such information can aid in developing remote patient monitoring systems that not only may inform parents (or other interested parties, including clinicians) about incidents of aggression in real-time but also potentially predict these incidents. This would allow a more timely intervention to prevent escalation and adjust treatment.

Although physical activity was the primary focus of our study, demographic and clinical features also played a prominent role in identifying aggression epochs. Our findings showed that age and CBCL Total Score were the first and third most significant feature in our model, respectively. Although the training set of physical aggression epochs and non-aggression epochs were randomly selected from the data pool of all aggression and non-aggression epochs, the average age and CBCL Total Score for selected physical aggression epochs were 8.5 ± 1.3 years and 43.6 ± 29.5 , respectively. In contrast, these numbers were 9.5 ± 1.7 years and 28.2 ± 49.0 for the randomly selected non-aggression epochs, indicating a significant difference between physical aggression and non-aggression epochs ($p < 0.0001$). Younger age was the strongest feature in the model associated with aggression epochs. This was consistent with well-established findings of reduced aggression with development and

maturation of inhibitory and self-regulation processes [47–50]. These two objective features are relatively easy to document reliably and appear to substantially contribute to a model.

Of the top four features contributing to the model's performance, only the CBCL score is based on subjective report. As would be expected, a higher score (denoting more severe psychiatric difficulties) is associated with a greater likelihood of physical aggressive incidents [51]. The CBCL is based on parental report and has excellent psychometric properties. It is a "broad-band" measure that captures various aspects of psychopathology, over the prior 6-month period. As such, it is quite distinct from many subjective measures of aggression, that are more directly temporally linked to such incidents. In other words, while the CBCL has the shortcoming of being subjective, it provides the necessary context within which to "interpret" an epoch as aggression or non-aggression.

In our study, of the 20 features considered in our model, sex had the least impact on its performance. This might appear surprising at first glance, given the well-established preponderance of aggression among boys. However, it is likely that the effect of sex was mediated by other variable including height, acceleration, overall activity level, etc.

We observed that the majority of physical aggression episodes occurred during the afternoon and evening, and on weekend days. This finding was consistent with previous research indicating that physical aggression among preschool-aged children is more likely to occur in the late afternoon and early evening hours and on weekends compared with weekdays [49]. One potential explanation is that these are the times when the children are with their parents or caregivers who are serving as the informants. This underscores the importance of optimizing the process to capture problematic behavior by making it less informant-dependent. Other potential factors explaining the higher likelihood of physical aggression during the afternoon, evening, and weekends include poorly controlled ADHD symptoms, increased free time, reduced supervision, social and peer pressure, and fatigue or stress [52,53].

It is important to note that the findings of this study do not suggest that children who exhibit the identified features are more likely to have aggressive incidents. Rather, these features may serve as potential indicators for identifying such episodes. To establish a causal association between these features and the prevalence of aggressive incidents, interventional studies are needed. Potential applications of these findings in clinical practice may include exporting physical activity data from smartphones or wearable sensors during clinical visits, with a particular focus on technology that exports vector magnitude. Additionally, a clinical tool could be developed to remotely track aggressive incidents, especially during the school day when parents/caregivers may not be aware of their occurrence.

While to our knowledge this is the first study to examine aggressive incidents in children in a real-life setting, several limitations should be acknowledged. Results may be skewed by a few participants with significantly longer time periods of physical aggression compared with other participants. To address this limitation, we performed a secondary analysis and down-sampled the number of epochs from participants who had considerably longer durations of physical aggression epochs compared with other participants. We identified one participant with a very long duration of a physically aggressive event (45 min). When the number of epochs for this subject was down-sampled and the model was fitted again, the model's performance was minimally impacted (precision = $76.4 \pm 5.5\%$ [decreased by 3.8%], accuracy = $78.6 \pm 4.9\%$ [decreased by 3.4%], recall = $83.9 \pm 8.4\%$ [decreased by 1.1%], F1 score = $79.7 \pm 4.9\%$ [decreased by 2.7%], and area under the curve = $85.7 \pm 5.4\%$ [decreased by 3.6%]). Nonetheless, our findings should be confirmed in a larger sample.

ADHD is a well-known risk factor for aggression but neither the presence of ADHD nor of aggression was required for study participation. As such, future studies should include children exhibiting higher aggression severity to confirm the generalizability of our findings. In the current study, similar to others [39,40], our approach used a random forest classifier to construct a classification model. Additionally, we conducted a secondary analysis to

demonstrate the relative importance of each feature. Although we assessed the optimized feature selection through cross-validation and bootstrapping, further validation using an independent dataset is necessary to confirm our findings and ensure their robustness and generalizability. Moreover, future work with larger sample sizes must explore other machine learning techniques and approaches, including innovative models seeking to better capture complex human behavior [54]. Another limitation is that aggressive incidents were recorded by parents/caregivers using a paper log, with instructions to record the events promptly. Nonetheless, their report may have been subject to recall bias or inaccuracy. To address this limitation, future studies could include teacher informants and/or an online platform to immediately report aggressive incidents in a more timely and accurate manner. Moreover, ensuring consistent sensor use among children displaying aggressive behavior may be challenging. Practical limitations, including restricted battery life and discomfort from extended wear, also pose difficulties. To bolster the effectiveness and broad adoption of our approach, future research should investigate complementary solutions such as video tracking, e-tattoo sensors, textile sensors integrated into clothing (e.g., smart t-shirts), and alternative sensor designs such as pendant sensors, which might be more acceptable to the target population. Future improvement should also include providing real-life alerts to caregivers when an aggressive episode occurs. Our study could not differentiate verbal aggressive behavior. Although most verbal aggression may be accompanied by physical agitation, relying solely on accelerometer sensors may not be adequate for identifying such behavior. Future research should explore the use of complementary sensors capable of capturing voice and volume to better detect verbal aggressive behavior. Finally, future studies could develop a mathematical model to detect aggressive behavior with a higher accuracy and determine the relationship between time of medication use and aggressive incidents reported in the evening.

5. Conclusions

This study presents a promising approach to identify incidents of physical aggression in children, using physical activity monitoring and machine learning. Faster movement speed, measured by vector magnitude, in combination with age, height, and CBCL Total Score could help identify physically aggressive incidents. Further research is needed to confirm the validity of these findings and explore clinical applications, including for prevention or treatment adjustment. For example, identifying predictors of aggression may help in the timely deployment of de-escalation procedures. Finally, given known racial/ethnic inequities in responding to problematic behavior [55], a tool to capture aggression objectively may reduce structural racism in healthcare and educational settings.

Author Contributions: Data processing, statistical analysis, visualization, interpreting results, and writing—original draft preparation, C.P.; interpreting results, and writing—original draft preparation, M.G.F.; machine learning model design and run, results visualization, M.D.R. and M.M.U.A.; study design and interpreting results, B.N. and C.C.; writing—review and editing, all authors; patient referral and screening, C.C., G.B.-V., E.C., C.A. and J.N.; technology development, coding, and technical support, C.P., M.M.U.A., R.K.M. and M.D.R.; resources, supervision, and project administration, C.C. and B.N.; funding acquisition, C.C. and B.N. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded in part by a grant from National Institutes of Health, NIH (award number: R01HD101326) and National Science Foundation Center to Stream Health in Place, C2SHIP (award number: NSF 2052514). The funding sources had no role in the study design, methods, data collection and analysis, and submission of the results.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Baylor College of Medicine (protocol code: H-47530; date of approval: 9 June 2020).

Informed Consent Statement: Informed consent was obtained from the parents/guardians of all participants enrolled in the study and assent was obtained from the participants.

Data Availability Statement: The de-identified datasets (except sensor data) are available upon request to the corresponding author(s).

Acknowledgments: We thank Maria Noun for supporting the coordination of some of the activities, and all participants and their parents who supported this study.

Conflicts of Interest: R.M. is now with BioSensics LLC., and M.A. is now with Abbott. However, their contributions to this study were limited to when they were postdoctoral associates with Baylor College of Medicine. The other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Gerson, R.; Malas, N.; Feuer, V.; Silver, G.H.; Prasad, R.; Mroczkowski, M.M. Best Practices for Evaluation and Treatment of Agitated Children and Adolescents (BETA) in the Emergency Department: Consensus Statement of the American Association for Emergency Psychiatry. *West. J. Emerg. Med.* **2019**, *20*, 409–418. [\[CrossRef\]](#)
2. Oerbeck, B.; Overgaard, K.R.; Aspenes, S.T.; Pripp, A.H.; Mordre, M.; Aase, H.; Reichborn-Kjennerud, T.; Zeiner, P. ADHD, comorbid disorders and psychosocial functioning: How representative is a child cohort study? Findings from a national patient registry. *BMC Psychiatry* **2017**, *17*, 23. [\[CrossRef\]](#)
3. Swanson, J.M.; Sergeant, J.A.; Taylor, E.; Sonuga-Barke, E.J.; Jensen, P.S.; Cantwell, D.P. Attention-deficit hyperactivity disorder and hyperkinetic disorder. *Lancet* **1998**, *351*, 429–433. [\[CrossRef\]](#)
4. Steinhausen, H.C.; Nøvik, T.S.; Baldursson, G.; Curatolo, P.; Lorenzo, M.J.; Rodrigues Pereira, R.; Ralston, S.J.; Rothenberger, A. Co-existing psychiatric problems in ADHD in the ADORE cohort. *Eur. Child Adolesc. Psychiatry* **2006**, *15* (Suppl. S1), I25–I29. [\[CrossRef\]](#)
5. Jensen, C.M.; Steinhausen, H.C. Comorbid mental disorders in children and adolescents with attention-deficit/hyperactivity disorder in a large nationwide study. *Atten. Deficit Hyperact. Disord.* **2015**, *7*, 27–38. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Pliszka, S. Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **2007**, *46*, 894–921. [\[CrossRef\]](#)
7. Willcutt, E.G. The prevalence of DSM-IV attention-deficit/hyperactivity disorder: A meta-analytic review. *Neurother. J. Am. Soc. Exp. Neurotherapeutics* **2012**, *9*, 490–499. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Kraut, A.A.; Langner, I.; Lindemann, C.; Banaschewski, T.; Petermann, U.; Petermann, F.; Mikolajczyk, R.T.; Garbe, E. Comorbidities in ADHD children treated with methylphenidate: A database study. *BMC Psychiatry* **2013**, *13*, 11. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Mahone, E.M.; Denckla, M.B. Attention-Deficit/Hyperactivity Disorder: A Historical Neuropsychological Perspective. *J. Int. Neuropsychol. Soc. JINS* **2017**, *23*, 916–929. [\[CrossRef\]](#)
10. Foster, E.M.; Jones, D.E. The high costs of aggression: Public expenditures resulting from conduct disorder. *Am. J. Public Health* **2005**, *95*, 1767–1772. [\[CrossRef\]](#)
11. National Institutes of Health Consensus Development Conference Statement: Diagnosis and treatment of attention-deficit/hyperactivity disorder (ADHD). *J. Am. Acad. Child Adolesc. Psychiatry* **2000**, *39*, 182–193. [\[CrossRef\]](#)
12. Mroczkowski, M.M.; Walkup, J.T.; Appelbaum, P.S. Assessing Violence Risk in Adolescents in the Pediatric Emergency Department: Systematic Review and Clinical Guidance. *West. J. Emerg. Med.* **2021**, *22*, 533–542. [\[CrossRef\]](#)
13. Mistler, L.A.; Friedman, M.J. Instruments for Measuring Violence on Acute Inpatient Psychiatric Units: Review and Recommendations. *Psychiatr. Serv.* **2022**, *73*, 650–657. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Connor, D.F.; Newcorn, J.H.; Saylor, K.E.; Amann, B.H.; Scahill, L.; Robb, A.S.; Jensen, P.S.; Vitiello, B.; Findling, R.L.; Buitelaar, J.K. Maladaptive Aggression: With a Focus on Impulsive Aggression in Children and Adolescents. *J. Child Adolesc. Psychopharmacol.* **2019**, *29*, 576–591. [\[CrossRef\]](#)
15. Althoff, R.R.; Ametti, M. Measurement of Dysregulation in Children and Adolescents. *Child Adolesc. Psychiatr. Clin. N. Am.* **2021**, *30*, 321–333. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Balia, C.; Carucci, S.; Coghill, D.; Zuddas, A. The pharmacological treatment of aggression in children and adolescents with conduct disorder. Do callous-unemotional traits modulate the efficacy of medication? *Neurosci. Biobehav. Rev.* **2018**, *91*, 218–238. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Pisano, S.; Masi, G. Recommendations for the pharmacological management of irritability and aggression in conduct disorder patients. *Expert Opin. Pharmacother.* **2020**, *21*, 5–7. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Findling, R.L.; Townsend, L.; Brown, N.V.; Arnold, L.E.; Gadow, K.D.; Kolko, D.J.; McNamara, N.K.; Gary, D.S.; Kaplin, D.B.; Farmer, C.A.; et al. The Treatment of Severe Childhood Aggression Study: 12 Weeks of Extended, Blinded Treatment in Clinical Responders. *J. Child Adolesc. Psychopharmacol.* **2017**, *27*, 52–65. [\[CrossRef\]](#)
19. Breaux, R.; Baweja, R.; Eadeh, H.-M.; Shroff, D.M.; Cash, A.R.; Swanson, C.S.; Knehans, A.; Waxmonsky, J.G. Systematic Review and Meta-analysis: Pharmacological and Nonpharmacological Interventions for Persistent Nonepisodic Irritability. *J. Am. Acad. Child Adolesc. Psychiatry* **2023**, *62*, 318–334. [\[CrossRef\]](#)
20. Kam, H.J.; Lee, K.; Cho, S.M.; Shin, Y.M.; Park, R.W. High-Resolution Actigraphic Analysis of ADHD: A Wide Range of Movement Variability Observation in Three School Courses—A Pilot Study. *Healthc. Inform. Res.* **2011**, *17*, 29–37. [\[CrossRef\]](#) [\[PubMed\]](#)

21. Martín-Martínez, D.; Casaseca-de-la-Higuera, P.; Alberola-López, S.; Andrés-de-Llano, J.; López-Villalobos, J.A.; Ardura-Fernández, J.; Alberola-López, C. Nonlinear analysis of actigraphic signals for the assessment of the attention-deficit/hyperactivity disorder (ADHD). *Med. Eng. Phys.* **2012**, *34*, 1317–1329. [\[CrossRef\]](#)
22. O'Mahony, N.; Florentino-Liano, B.; Carballo, J.J.; Baca-García, E.; Rodríguez, A.A. Objective diagnosis of ADHD using IMUs. *Med. Eng. Phys.* **2014**, *36*, 922–926. [\[CrossRef\]](#)
23. Jeannet, P.Y.; Aminian, K.; Bloetzer, C.; Najafi, B.; Paraschiv-Ionescu, A. Continuous monitoring and quantification of multiple parameters of daily physical activity in ambulatory Duchenne muscular dystrophy patients. *Eur. J. Paediatr. Neurol.* **2011**, *15*, 40–47. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Moreno, J.P.; Razjouyan, J.; Lester, H.; Dadabhoy, H.; Amirmazaheri, M.; Reesor-Oyer, L.; O'Connor, T.M.; Hernandez, D.C.; Najafi, B.; Alfano, C.A.; et al. Later sleep timing predicts accelerated summer weight gain among elementary school children: A prospective observational study. *Int. J. Behav. Nutr. Phys. Act* **2021**, *18*, 94. [\[CrossRef\]](#)
25. Welch, V.; Wy, T.J.; Ligezka, A.; Hassett, L.C.; Croarkin, P.E.; Athreya, A.P.; Romanowicz, M. Use of mobile and wearable artificial intelligence in child and adolescent psychiatry: Scoping review. *J. Med. Internet Res.* **2022**, *24*, e33560. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Lindhiem, O.; Goel, M.; Shaaban, S.; Mak, K.J.; Chikersal, P.; Feldman, J.; Harris, J.L. Objective measurement of hyperactivity using mobile sensing and machine learning: Pilot study. *JMIR Form. Res.* **2022**, *6*, e35803. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Goodwin, M.S.; Mazefsky, C.A.; Ioannidis, S.; Erdogmus, D.; Siegel, M. Predicting aggression to others in youth with autism using a wearable biosensor. *Autism Res. Off. J. Int. Soc. Autism Res.* **2019**, *12*, 1286–1296. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Achenbach, T.M.; Rescorla, L. *Manual for the ASEBA School-Age Forms & Profiles: An Integrated System of Multi-Informant Assessment*; ASEBA: Burlington, VT, USA, 2001.
29. Guerrero, S.; Menghini, D.; Napoli, E.; Di Vara, S.; Valeri, G.; Vicari, S. Assessment of Psychopathological Comorbidities in Children and Adolescents with Autism Spectrum Disorder Using the Child Behavior Checklist. *Front. Psychiatry* **2019**, *10*, 535. [\[CrossRef\]](#)
30. Busner, J.; Targum, S.D. The clinical global impressions scale: Applying a research tool in clinical practice. *Psychiatry* **2007**, *4*, 28–37.
31. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*; American Psychiatric Association: Arlington, VA, USA, 2013.
32. Ogden, C.L.; Kuczmarski, R.J.; Flegal, K.M.; Mei, Z.; Guo, S.; Wei, R.; Grummer-Strawn, L.M.; Curtin, L.R.; Roche, A.F.; Johnson, C.L. Centers for Disease Control and Prevention 2000 growth charts for the United States: Improvements to the 1977 National Center for Health Statistics version. *Pediatrics* **2002**, *109*, 45–60. [\[CrossRef\]](#)
33. Calarge, C.A.; Schlechte, J.A.; Burns, T.L.; Zemel, B.S. The effect of psychostimulants on skeletal health in boys co-treated with risperidone. *J. Pediatr.* **2015**, *166*, 1449–1454 e1441. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Blader, J.C.; Pliszka, S.R.; Kafantaris, V.; Foley, C.A.; Carlson, G.A.; Crowell, J.A.; Bailey, B.Y.; Sauder, C.; Daviss, W.B.; Sinha, C.; et al. Stepped Treatment for Attention-Deficit/Hyperactivity Disorder and Aggressive Behavior: A Randomized, Controlled Trial of Adjunctive Risperidone, Divalproex Sodium, or Placebo After Stimulant Medication Optimization. *J. Am. Acad. Child Adolesc. Psychiatry* **2021**, *60*, 236–251. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Leeger-Aschmann, C.S.; Schmutz, E.A.; Zysset, A.E.; Kakebeeke, T.H.; Messerli-Bürge, N.; Stülz, K.; Arhab, A.; Meyer, A.H.; Munsch, S.; Jenni, O.G.; et al. Accelerometer-derived physical activity estimation in preschoolers—Comparison of cut-point sets incorporating the vector magnitude vs the vertical axis. *BMC Public Health* **2019**, *19*, 513. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Howe, C.A.; Clevenger, K.A.; Leslie, R.E.; Ragan, M.A. Comparison of Accelerometer-Based Cut-Points for Children's Physical Activity: Counts vs. Steps. *Children* **2018**, *5*, 105. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Actigraph Documentation. Vector Magnitude. Available online: <https://actigraphcorp.my.site.com/support/s/article/What-is-VM-Vector-Magnitude> (accessed on 16 January 2023).
38. Actigraph Documentation. Kilocalories Equation. Available online: <https://actigraphcorp.my.site.com/support/s/article/What-is-the-difference-among-the-Energy-Expenditure-Algorithms> (accessed on 16 January 2023).
39. Rogers, J.; Gunn, S. Identifying Feature Relevance Using a Random Forest. In *Subspace, Latent Structure and Feature Selection*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 173–184.
40. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [\[CrossRef\]](#)
41. Japkowicz, N.; Stephen, S. The Class Imbalance Problem: A Systematic Study. *Intell. Data Anal.* **2002**, *6*, 429–449. [\[CrossRef\]](#)
42. Wang, Z.; Jiang, C.; Ding, Y.; Lyu, X.; Liu, Y. A Novel behavioral scoring model for estimating probability of default over time in peer-to-peer lending. *Electron. Commer. Res. Appl.* **2018**, *27*, 74–82. [\[CrossRef\]](#)
43. Sage, A.J. Random Forest Robustness, Variable Importance, and Tree Aggregation. Ph.D. Dissertation, Department of Statistics, Iowa State University, Ames, IA, USA, 2018.
44. Javeed, A.; Zhou, S.; Yongjian, L.; Qasim, I.; Noor, A.; Nour, R. An Intelligent Learning System Based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection. *IEEE Access* **2019**, *7*, 180235–180243. [\[CrossRef\]](#)
45. Ramadhan, M.M.; Sitanggang, I.S.; Nasution, F.R.; Ghifari, A. Parameter Tuning in Random Forest Based on Grid Search Method for Gender Classification Based on Voice Frequency. *DEStech Trans. Comput. Sci. Eng.* **2017**, *10*. [\[CrossRef\]](#)
46. Adelabu, S.; Mutanga, O.; Adam, E. Testing the reliability and stability of the internal accuracy assessment of random forest for classifying tree defoliation levels using different validation methods. *Geocarto Int.* **2015**, *30*, 810–821. [\[CrossRef\]](#)

47. Lee, K.H.; Baillargeon, R.H.; Vermunt, J.K.; Wu, H.X.; Tremblay, R.E. Age differences in the prevalence of physical aggression among 5-11-year-old Canadian boys and girls. *Aggress Behav.* **2007**, *33*, 26–37. [[CrossRef](#)] [[PubMed](#)]
48. Teymoori, A.; Cote, S.M.; Jones, B.L.; Nagin, D.S.; Boivin, M.; Vitaro, F.; Orri, M.; Tremblay, R.E. Risk Factors Associated with Boys' and Girls' Developmental Trajectories of Physical Aggression From Early Childhood Through Early Adolescence. *JAMA Netw. Open* **2018**, *1*, e186364. [[CrossRef](#)] [[PubMed](#)]
49. Broidy, L.M.; Nagin, D.S.; Tremblay, R.E.; Bates, J.E.; Brame, B.; Dodge, K.A.; Fergusson, D.; Horwood, J.L.; Loeber, R.; Laird, R.; et al. Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Dev. Psychol.* **2003**, *39*, 222–245. [[CrossRef](#)] [[PubMed](#)]
50. Tremblay, R.E. Early development of physical aggression and early risk factors for chronic physical aggression in humans. *Curr. Top. Behav. Neurosci.* **2014**, *17*, 315–327. [[CrossRef](#)] [[PubMed](#)]
51. Navarro, R.; Larranaga, E.; Yubero, S.; Villora, B. Families, Parenting and Aggressive Preschoolers: A Scoping Review of Studies Examining Family Variables Related to Preschool Aggression. *Int. J. Environ. Res. Public Health* **2022**, *19*, 5556. [[CrossRef](#)] [[PubMed](#)]
52. DuRant, R.H.; Getts, A.; Cadenhead, C.; Emans, S.J.; Woods, E.R. Exposure to violence and victimization and depression, hopelessness, and purpose in life among adolescents living in and around public housing. *J. Dev. Behav. Pediatr.* **1995**, *16*, 233–237. [[CrossRef](#)]
53. DuRant, R.H.; Cadenhead, C.; Pendergrast, R.A.; Slavens, G.; Linder, C.W. Factors associated with the use of violence among urban black adolescents. *Am. J. Public Health* **1994**, *84*, 612–617. [[CrossRef](#)]
54. Kumar Jain, D.; Liu, X.; Neelakandan, S.; Prakash, M. Modeling of human action recognition using hyperparameter tuned deep learning model. *J. Electron. Imaging* **2022**, *32*, 011211. [[CrossRef](#)]
55. Compliance Review. Available online: <https://www2.ed.gov/about/offices/list/ocr/docs/investigations/03105001-a.html> (accessed on 11 April 2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.