The Intensity Inequality index for physical activity: a new metric for integrative analysis of movement

Abstract

Background: Wearable sensors recording acceleration provide a powerful tool for analysis of physical activity (PA). Continuous, high-rate data acquisition over extended periods gives highly resolved measurement of movement intensity. Whilst increased complexity of PA analytics allows for deeper insight it brings a challenge to statistical testing, where commonly used approaches require a single defining metric for PA per participant. ***Methods:*** We adapt an Econometrics measure to obtain a statistical test metric for movement intensity - the *intensity inequality index*, $I_{\neq}$. This is a 'Gini coefficient for movement' that quantifies the inequality in distribution of time spent across a range of activity intensity values. The $I_{\neq}$ metric is calculated using a graphical method on plots of cumulative time versus cumulative intensity level. Hypothesis testing of $I_{\neq}$ is performed on 24-hour activity traces of 58 children, aged 7-11 years, to assess statistical differences in PA between typically developing children (TD) and those suspected of having developmental coordination disorder (sDCD). ***Results:*** The $I_{\neq}$ test metric provided high statistical confidence with low sample numbers: p-value < 0.05 for n ≥ 30. When differentiating between groups $I_{\neq}$ halved the sample size required for a statistical power of 80% at α = 0.05, in comparison to the alternative metrics of intensity gradient or log ratio of minutes at low and moderate-high intensity. ***Conclusions:*** The inequality index provides a metric that is based on the accumulated time-counts across an activity intensity distribution. This integrative description of the distribution makes it a powerful statistical metric for PA.

Background

The quantification of physical activity (PA), both in duration, intensity and volume, is essential in assessing the relationship and impact of movement behaviour on health outcomes.[1,2] To this end, acceleration is the predominant measurement variable of choice[3,4] and is routinely measured using electronic sensors (accelerometers).[5] The translation of acceleration into PA usually involves combination of tri-axial vectors into a single normalised summary metric, e.g. mean absolute deviation (MAD), monitor independent movement summary (MIMS) or Euclidean norm minus one (ENMO),[6] together with time integration over a defined epoch (generally 1-60 seconds) to provide an activity volume metric such as steps,[7,8] or classification into broad activity levels using pre-determined acceleration cut-point values (i.e. sedentary, light, moderate-to-vigorous).[9] Whilst the discrimination of active periods from sedentary behaviour is epidemiologically important,[10-12] utilising richer metrics beyond cut-points can further differentiate volume and intensity of activity,[13-15] and describe the full spectrum of measured intensity and distribution of daily activity in more detail.[16,17] This has been aided by the improvement of wearable, low-cost sensors capable of measuring acceleration with high temporal resolution (100 Hz) and with sufficient on-board memory to store data from a full week of measurement. While these sensors continuously monitor PA intensity throughout the day, there is still significant potential in developing metrics that fully reflect the power of this metrology to improve the quality, efficiency and interpretation of surveillance, observation and intervention studies in the field.

A full description of the range of exertion during PA is provided by frequency distributions of time spent across quasi-continuous definitions of intensity level[18], sometimes referred to as the PA intensity spectrum. This is quantified by the intensity gradient, I_{grad}: a single metric corresponding to the slope of the regression line to the distribution.[16,18] Given the opportunity for the development of new analytics in this field of research we review alternative measures to I_{grad}, which may provide greater

power when testing for statistical differences between groups or across interventions. There is a significant opportunity to capture the heterogeneity in PA intensity with a single integrative measure – i.e. to find a quantifier of the difference between frequency distributions. While this is a new concept in PA research, Econometrics, has a rich literature on the quantification of wealth distributions across a population.[19-21] A range of statistical metrics based on variance,[20] rank ordering,[20] entropy,[22] or obtained from graphical constructions[19] have been widely used. These wealth inequality metrics describe resource allocation, where the resource is the monetary value of wealth, and the distribution considered is that across a given human population. The key step change in the context of PA is that the same metrics can be used, with minutes-spent being the resource that is allocated across an activity intensity (ENMO) range.

The use of a metric based on the PA intensity spectrum undoubtedly adds complexity to data calculation and interpretation in comparison to averaged metrics or activity cut-points. However, whilst the methods will seem unfamiliar to some in the field, they are standard practice in other scientific disciplines, and the benefits of using a more nuanced descriptor of this type for PA analysis have already been demonstrated in the intensity gradient metric.[16,17] The central contribution of this work lies in its ability to harness significantly more data than conventional approaches, thereby increasing statistical power. This addresses a persistent limitation in device-based physical activity measurement: high variance caused by excessive data reduction. By minimizing unnecessary reduction, our approach enables more precise estimation with smaller sample sizes. This has meaningful implications for the design of intervention and cohort studies, potentially reducing both sample size requirements and associated funding needs. Given the high costs of physical activity measurement in large-scale studies, we believe our method provides a cost-effective and scalable alternative. This could be particularly valuable in resource-constrained contexts and for funding bodies seeking efficiency without compromising data quality.

A framework for analysing the 'equality' of distribution of a variate across a value range was presented by Dalton, over a century ago.[23] This is based on a concept of equality of allocation across the frequency distribution, measured by an inequality metric which is equal to 1 for a uniform distribution (constant count frequency across the range) and equal to 0 for maximal concentration of resource into a single allocation unit (e.g. all wealth held by a single individual). Dalton uses the compositional nature of a fixed-resource allocation[24,25] to construct a model with which to compare distributions. A reduced allocation of resource to 1 level must mean an increase at another, this concept of a 'transfer' of resource can be used to assess the capability of a proposed metric to capture differences between distributions.[26] For example, a simple measure of resource allocated to different ranges, such as quartiles, quantifies different dispersions but does not provide a unique value – local transfers between close-valued distribution bins change the distribution but may leave the quartile values unaltered. For the same reason, metrics of PA based on time spent in different intensity categories (e.g. log ratios of minutes at different levels)[27] will not resolve differences due to reallocation of time within a category, e.g. transfer of time from moderate to vigorous intensity. The standard deviation of the distribution metric quantifies differences at individual bin level but is limited to using a common reference point - the mean value. A more discriminatory metric is provided by the Gini coefficient,[28] this is based on cumulative distributions which incorporate information from across the distribution through summation of frequency values. This provides the Gini coefficient with high statistical power and it is the most widely used measure of wealth inequality.[29] It is commonly applied to wealth distributions across a human population but has also been applied to PA research to quantify the distribution of activity metrics across a population.[30,31] In this paper we adapt the mathematical formulism of the Gini coefficient in a different way, applying it to the distribution of time across PA intensity level, that is, it quantifies the PA intensity distribution from an individual rather than the activity distribution across a cohort. We define an inequality index of PA intensity – referred to as $\boldsymbol{I}_{\neq}$, which is based on cumulative summations of minutes spent at different activity intensity levels and

thus introduces a novel statistical metric for developing research designs and new insights in exercise science.

Within the collection of potential descriptors of PA, I_{grad} and $I_{\neq}$ are similar in that they both seek to provide a metric that describes the intensity distribution of activity. We therefore briefly compare and contrast the 2 metrics. Both use the distribution of minutes spent across a range of intensity values; I_{grad} is based on data regression and formulated as the slope of the regression line, $I_{\neq}$ is based on summation of data values and formulated as the inequality of the minute counts relative to a uniform distribution. I_{grad} may be limited due to the assumption of linearity in the log-log data, thus the R^2 regression value must be considered alongside the value of the metric; the main limitation to $I_{\neq}$ is that it is not unique – the unevenness of counts is quantified rather than the absolute position within the distribution, different distributions may therefore give identical inequality index values. Both metrics have an immediate interpretation – changes in the slope of the distribution indicates movement of time spent from high/low to low/high intensity; greater inequality stems from concentration of time into limited spans of intensity levels. $I_{\neq}$ also provides a bounded numerical range with clear limiting values, 0 → equal time spent at all intensities, 1→ all time spent at one intensity. Finally, $I_{\neq}$ clearly differs from I_{grad} in that it makes no assumption as to the form of the data and can calculate the cumulative time-spent from any given intensity distribution. In application both metrics provide additional discrimination over simple activity means or blunt categorisation into a few activity levels. Their relative performance will be dependent on the specifics of the given dataset and so $I_{\neq}$ is best viewed as an additional statistical tool that sits alongside I_{grad} , with the choice between the 2 metrics needing to be based on trial application rather than decided a-priori.

The aim in this paper is to present the new metric, $I_{\neq}$ and demonstrate its validity for PA analysis through application to data acquired in a study by Swindell et al of motor competence in school-age

children.[32] The hypothesis being that greater statistical discrimination is possible due to the integration of PA intensity information within this intensity inequality index. Motor competence is recognised as an important enabler of PA and several systematic reviews have demonstrated a positive association between motor competence and PA.[33,34] Recent research has documented high rates of poor motor competence (30–77%) in school-age children, with 5–6% of this group diagnosed as having Developmental Coordination Disorder (DCD) – a neurodevelopmental disorder characterised by a marked impairment of motor competence.[35] Swindell's study[32] provides wide heterogeneity between participants in the intensity distribution of PA, thus is useful to test new analytics in PA research. In this study, we use $I_{\neq}$ as a test statistic to describe the active parts of 24-hour movement behaviours in children, classified into typically developing (TD) and suspected DCD (sDCD) groups, and compare it to alternative metrics.

Methods

Participants and settings

All data were collected from the Moves-UP project, a school-based intervention aimed at improving motor coordination in primary school children with suspected developmental coordination disorder (sDCD).[32] Only baseline data from the project was used for this study. The protocols for the project were approved by the Swansea University ethics committee (approval no: 3 2023 6474 6381) and written informed consent (Parents and schools) and assent (children) were obtained from all participants prior to data collection. Children from four schools in Swansea, participated in this study. Data was collected between the 1st of November 2023 and the 30th of January 2024 from a convenience sample that included 58 children (55% boys, aged 8.6 $\pm$1.6 years).

Measurements

Participants wore an Axivity AX3 (Axivity Ltd, Newcastle, UK) accelerometer on the non-dominant wrist 24 hours per day for 7 days with accelerations recorded at 100 Hz and a dynamic range of $\pm$ 8 g.[36] Data

were downloaded using OmGui open-source software (OmGui v 1.0.0.43, Open Movement, Newcastle University, UK). All data were processed in Matlab (MathWorks, Natick, MA, U.S.A.). Raw triaxial accelerations, ax, ay and az, were converted into a single, scalar measure of acceleration using the Euclidean Norm minus one (ENMO) with acceleration in mg units:

$$ENMO = \sqrt{ax^2 + ay^2 + az^2} - 1000 \qquad \text{[equation 1]}$$

ENMO values were averaged over 1-second epochs for further analysis.[6] Negative values were rounded to zero. The purpose of the study was to assess metrics that can quantify movement intensity, it did not aim to accurately classify periods of inactivity. A straightforward approach was therefore taken to data filtering that selected only data from a 24-hour period corresponding to movement. Non-active periods (sleep, inactivity or non-wear time) were treated as a single class, defined by a threshold: ENMO < 1 mg. The 1 mg threshold was chosen as a reasonable level based on calculation of ENMO during sleep periods of the 24-hour trace. We wish to remove periods with negligible acceleration (inactivity) but keep data at low acceleration (what we would refer to as sedentary time). Low-level acceleration that normally falls within a 'sedentary' classification was kept in the data trace. The mean inactive time per day was 1132±192 minutes and is comparable to that measured in similar studies[16]. Participants were included provided at least one full 24-hour period was recorded and all days on which the sensor was worn were included (weekdays and weekends). PA cut-points typical of reported values were used to define sedentary behaviour (SB), low intensity physical activity (LPA), and moderate to vigorous physical activity (MVPA), (12 mg ≤ SB ≤ 50 mg , 50 mg < LPA < 200 mg, 200 mg ≤ MVPA).[32, 37, 38]

Motor coordination

To assess potential risk of DCD, parents/carers of each participating child completed the Developmental Coordination Disorder Questionnaire (DCDQ+).[39] Children who scored below 56 (out of 75) were classified as sDCD for subsequent analysis. To mitigate against potential response bias,

respondents were not given information of survey threshold or outcomes. Furthermore, teachers provided support in completing the questionnaire to parents/guardians who needed it.

Statistical methods

Intensity gradient: The variation in activity level was visualised through the intensity distribution, this is a time histogram in which the 1440 minutes from a 24 hour period are binned according to ENMO level.[17,18] 100 bins were used with either a linear binning, $ENMO(n)_{1:100} = n \times 25$ mg, or a logarithmic binning, $ENMO(n)_{1:100} = 10^{n/30}$ mg (the rationale for this is explained in the Results section). The slope of the intensity gradient, I_{grad} was calculated using linear regression on the linearly binned, time-ENMO distribution.[18]

Inequality index: The intensity gradient is quantified by the *intensity inequality index*, $I_{\neq}$. This is obtained from cumulative distributions of activity time (minutes) and ENMO level (the bins of the intensity distribution). Graphically, this is akin to a P-P plot of cumulative distributions,[40] a specific instance of which – the Lorenz curve, is commonly used in economics to assess wealth distribution across a population.[19] The wealth inequalities described by a Lorenz curve are quantified by the Gini coefficient,[28] a metric that spans a range of 0 → 1, where a value of 1 indicates maximum inequality (all wealth held by a single individual) and a value of 0 indicates equal sharing of wealth between all individuals. The activity intensity inequality index, $I_{\neq}$, presented here adapts the formulism of the Gini coefficient, substituting allocation of activity minutes for allocation of wealth. The value of $I_{\neq}$ can be graphically obtained from areas in a P-P plot (*see figure 1 and accompanying text*),), It is mathematically defined as the difference in minutes accumulated across the measured intensity distribution from that which would accumulate with uniformly distributed time, summed over the range of activity intensities:

$$I_{\neq} = 2 \times \frac{\sum_{j=1}^{n}\left[\sum_{i=1}^{j} m_i - j\bar{m}\right]}{n} \qquad \text{[equation 2]}$$

Where m_i represents minutes spent at activity intensity level i, $\bar{m}$ is the average minutes per intensity level and n is the total number of bins (n = 100 for all calculations shown). Note: in this formulation the minute counts are calculated as fractional values, i.e. $\sum_{i=1}^{n} m_i = n\bar{m} = 1$. The intensity distribution for PA is naturally heavily biased to low ENMO values, to ensure good discrimination across the range of activity intensity we therefore use the $\log_{10}$ ENMO bin-values described above and express m_i in $\log_{10}$ units. As the $I_{\neq}$ index is based on cumulative fractional distributions, it remains valid under this logarithmic scaling.

Wilcoxon Rank testing: The $I_{\neq}$ index is non-normal and so the non-parametric Wilcoxon Rank test was used for statistical hypothesis testing. Random sampling from the set of 24-hour activity profiles (n = 336) was used to calculate statistical p-values. At each specified sample size, $N_{24\,hour}$, the Wilcoxon rank test was computed for 1,000 random selections of size $N_{24\,hour}$. The quoted p-value is the mean value from the 1,000 runs.

Statistical power: The non-normality of the inequality test statistic invalidates standard references or calculators for statistical power,[41] for this reason we adopted a Monte Carlo simulation approach.[42] The Wilcoxon rank test was performed 1,000 times for random selections of $N_{24\,hour}$ traces of 24-hour activity. Testing of sample pairs where both are activity profiles of typically developing children (TD) provides the distribution of the Wilcoxon test statistic where the null hypothesis holds true. Testing of TD activity profiles versus sDCD profiles, provides the distribution of the Wilcoxon test statistic where the null hypothesis is rejected. Comparison of the overlap of the two test statistic distributions provides Type I (α) and Type II (β) errors and the calculation of statistical power.

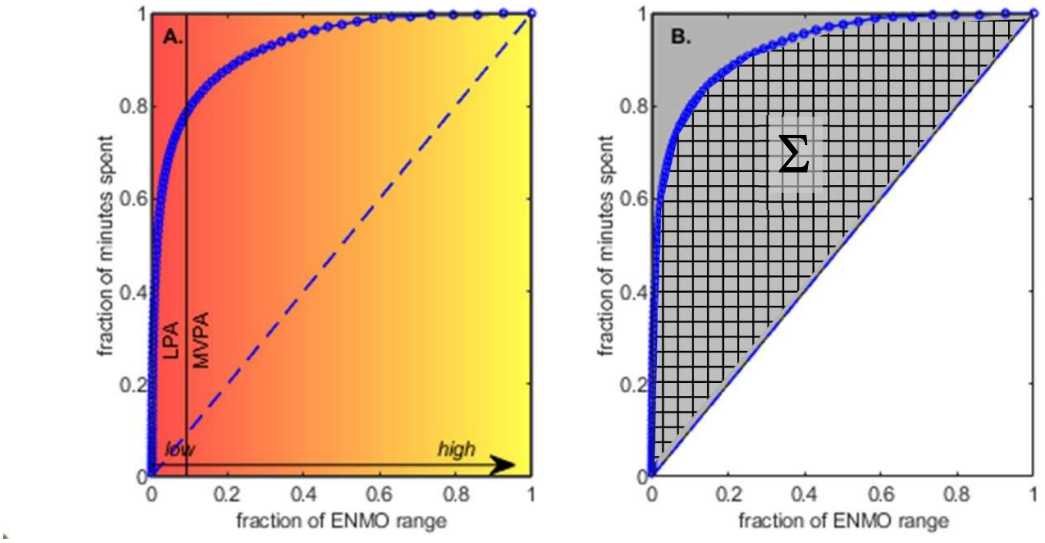

Figure 1 – (A) Cumulative time versus cumulative intensity for a typical 24-hour activity trace. (B) Graphical calculation of $I_{\neq}$ from area under the Lorenz curve. ENMO indicates Euclidean norm minus one; LPA, low-intensity physical activity; MVPA, moderate to vigorous physical activity

Results

Profiling the intensity of physical activity

A complete description of PA intensity and duration is provided by a time-intensity, frequency distribution.[17] As defined by Rowlands *et. al.* this is a distribution of time spent across activity levels, where the ENMO value of measured acceleration is the intensity metric. This is displayed on a ln-ln plot (example shown in *figure 2A*). The logarithmic axes ensure that the small fractions of time spent in high intensity activity are clearly visible.

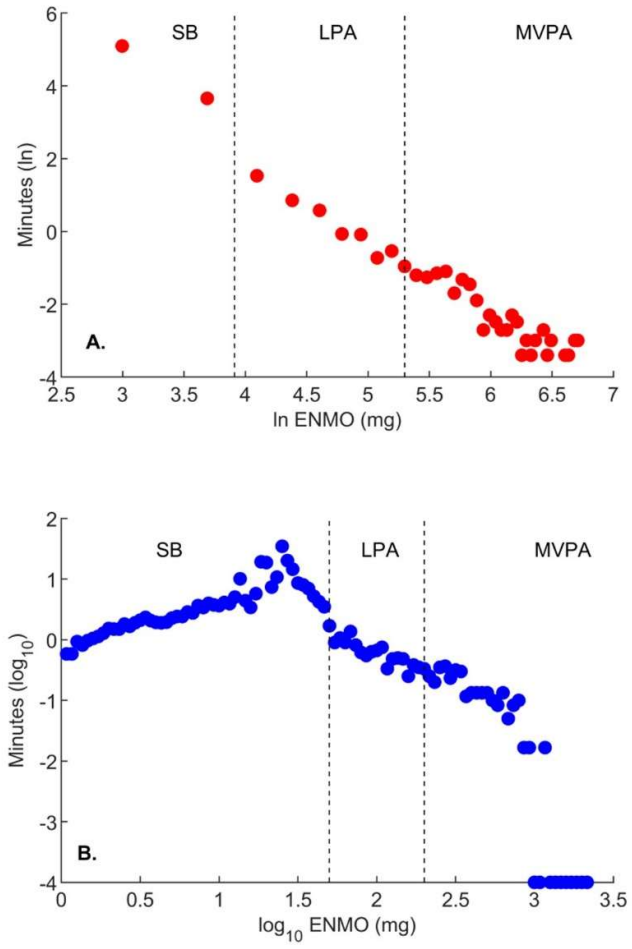

Figure 2 — Frequency distributions of time spent at different PA intensity. (A) Linear ENMO bins. (B) Log10 ENMO bins. ENMO indicates Euclidean norm minus one; LPA, low-intensity physical activity; MVPA, moderate to vigorous physical activity; PA, physical activity; SB, sedentary behavior.

Whilst the display uses a logarithmic intensity-axis, the distribution is formed by a linear time-binning (Δ_{ENMO} = 25 mg). However, the spread of time spent at different activity levels is far from linear as most of the day is spent in sedentary behaviour (SB) or light activity (LPA). The distribution of plot points is therefore highly skewed, i.e. in Fig. 2A, SB accounts for 98.9% of daily minutes but is represented by only 2 of 100 data points whereas MVPA is undertaken for only 0.003% of the time yet is displayed

using 90 of the 100 points. It is common practise when handling logarithmic data to use logarithmic bins when creating the data distribution to ensure equidistant points in a logarithmic plot.[43] We take this approach to produce an adapted intensity distribution (Fig. 2B) where minutes spent are binned into $\log_{10}$ intervals ($\Delta_{ENMO}(n)_{1:100} = 10^{n/30}$ mg, for intensity bins with t=0, the y-value is set to 10^{-4}). This produces a plot in which the low intensity activity is much better represented (SB is covered by 51 of 100 data points) whilst the high intensity activity remains visible. It is notable that the additional resolution at low intensity shows structure that is hidden in the standard intensity distribution. There is a clear peak in activity between 1-1.5 $\log_{10}$ units (ENMO ~12 – 30 mg). Thus, the finer discrimination of the log-binned distribution provides the resolution to interrogate low intensity activity and develop a more nuanced understanding of the range of physical movement that lies within the 'sedentary' categorisation.

The intensity inequality index

A typical plot of the cumulative time spent at different ENMO levels, for 24 hours of activity, is shown in figure 1A. To maintain good discrimination across all of the activity intensity range we use the log-transformed data values of minutes spent and ENMO (i.e. data in the format shown in figure 2B). To obtain a distribution metric we use the Lorenz curve to define an *Intensity Inequality* index in similar manner to the Gini coefficient, we represent this with the symbol: $\mathrm{I}_{\neq}$. This inequality index is defined graphically as the hatched area shown in figure 1B divided by the shaded triangular area, i.e. $\mathrm{I}_{\neq} = {\Sigma}/{0.5}$. The index tends to a limit of zero for equal time spent across all intensity levels (Lorenz curve follows the diagonal and $\Sigma \to 0$), or to one for total inactivity (ENMO = 0). Relating the graph to mathematical definition provided earlier (equation 2) – the points on the accumulation curve are given by the term, $\sum_{i=1}^{j} m_i$, the diagonal line by, $\sum_{i=1}^{j} i\overline{m}$ and the hatched area represents the outer summation from $j = 1\ to\ n$.

As $I_{\neq}$ is determined by the shape of the cumulative time distribution function (CDF) it is sensitive to all of the activity intensity data and so provides the integrative metric that we require to accurately quantify differences in intensity distribution profiles. To visually demonstrate this, two activity intensity Lorenz curves are plotted in figure 3. These represent 24 hours of activity by children at either end of the $I_{\neq}$ range. In comparing the two datasets the shaded areas show additional time spent in LPA by child a. and additional time engaged in the high end of MVPA by child b. This gives a visual illustration of the transfer concept outlined by Dalton[23] – in which the curves may be viewed as a record of the apportioning of the fixed time resource across PA levels. From this viewpoint, the greater activity intensity of child b. relative to child a., is achieved through a transfer of a block of LPA time into high intensity movement (as indicated by the arrow in figure 3).

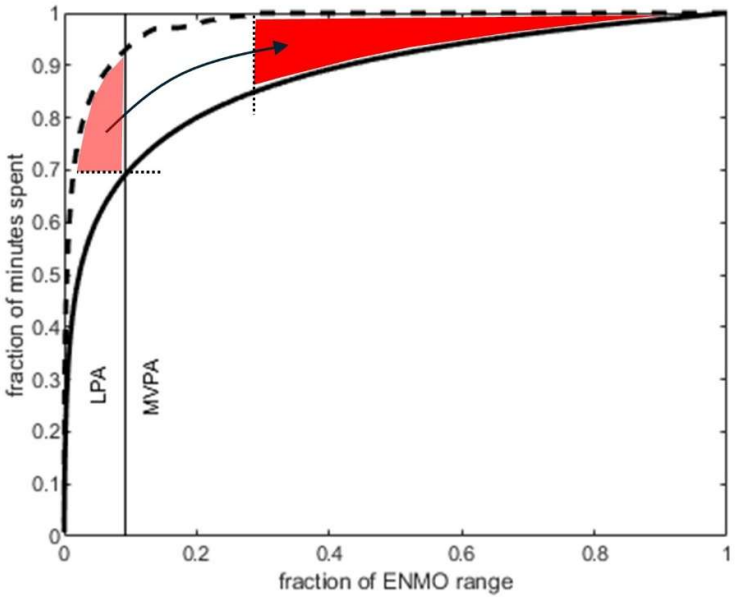

Figure 3 – Examples of time-ENMO Lorentz curves for high inequality (a. dashed line, $I_{\neq}$ = 0.96) and low inequality (b. solid line, $I_{\neq}$ = 0.76) distributions. Shaded areas highlight differences between data sets a and b for time spent in low- and high-intensity activity. ENMO indicates Euclidean norm minus one; LPA, low-intensity physical activity; MVPA, moderate to vigorous physical activity.

To further demonstrate the ability of $I_{\neq}$ to capture differences in activity profiles we consider the effect of the transfer of a set number of minutes from low intensity to high intensity activity (figure 4). Figure 4A. shows a simulated transfer of 10 minutes of activity from SB to MVPA classes, using the measured distribution shown in figure 1B. The cumulative time and activity intensity are shown in figure 4B. The key observation is that the integrative function of the cumulative distributions alters the curves for all values spanning the interval between the altered bin entries – i.e. the removal of minutes spent at low intensity reduces the cumulative time at this intensity level and this reduction remains, across the curves until it is corrected at high intensity by the addition of these minutes. Thus the $I_{\neq}$ metric is particularly sensitive to altered PA, where the change involves large differences in activity intensity level.

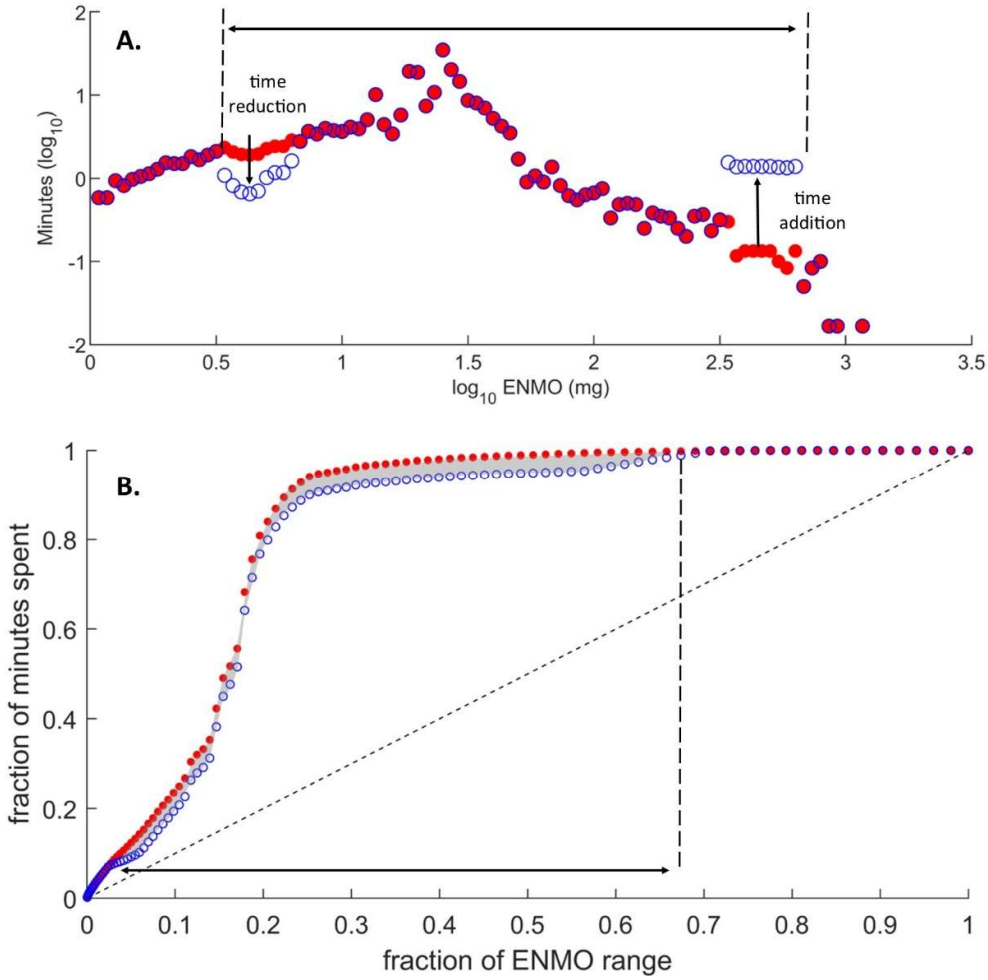

Figure 4 – (A) Frequency distributions of time spent at different PA intensity. Original trace: red (filled) circles, simulated transfer of 10 minutes of activity from low to high intensity, and blue (open) circles, 10 minutes divided equally across altered bins. (B) Cumulative time versus cumulative intensity for the distributions. (Note: arrows in figure depict the PA intensity span encompassed by the shifted time allocation). ENMO indicates Euclidean norm minus one.

Hypothesis testing using the inequality index

To demonstrate the use of $I_{\neq}$ in statistical testing we compare PA data collected from a cohort of 58 primary school children, a sub-set of which have sDCD. The distribution of $I_{\neq}$ values, from multiple periods of 24-hour activity, for 30 TD children is shown in figure 5A. The data is non-normal and so we

adopt the non-parametric, Wilcoxon rank-sum test for the statistical analysis. The distribution of the Wilcoxon test statistic comparing datasets selected at random from the same TD cohort follows a standard normal distribution (figure 5B), as expected for this case where the null hypothesis holds true.

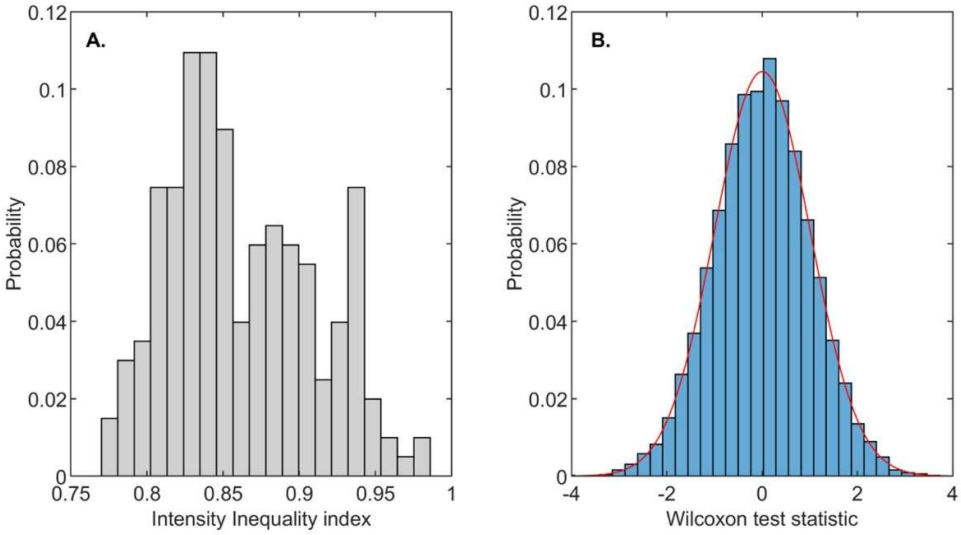

Figure 5 – (A) Probability distribution of $I_{\neq}$ for TD children (n_{child} = 30, $N_{24\ hour}$ = 200). (B) Probability distribution of Wilcoxon test statistic for comparison of $I_{\neq}$ for randomly selected TD cohorts ($N_{24\ hour}$ = 100, n_{trials} = 1000; solid line: standard normal distribution). TD indicates typically developing.

Previous work, using PA intensity cut-offs has shown differences between TD and sDCD groups, with increased time being spent in sedentary behaviour by the sDCD children.[32] We therefore use this as a benchmark from which we can assess the utility of $I_{\neq}$ as a test metric. A comparison of summary statistics for TD and sDCD groups obtained from a log ratio composition,[27] the intensity gradient and the $I_{\neq}$ index is shown in figure 6, together with conventional metrics of average acceleration and total activity minutes. Considering the metrics that quantify the intensity of activity: there is no statistical difference between the average acceleration of the two groups; however, the three other intensity

metrics all show a statistically significant difference between sDCD and TD groups ($p < 0.05$). The $I_{\neq}$ index provides the highest statistical confidence with greater separation of test metric distributions. The enhanced discrimination provided by the inequality metric is expected as this is an integrative index, encompassing the full range of activity intensity (100 bins of the intensity gradient). In contrast, a compositional ratio using activity cut-points can only capture two measurements on limited spans of the intensity range.

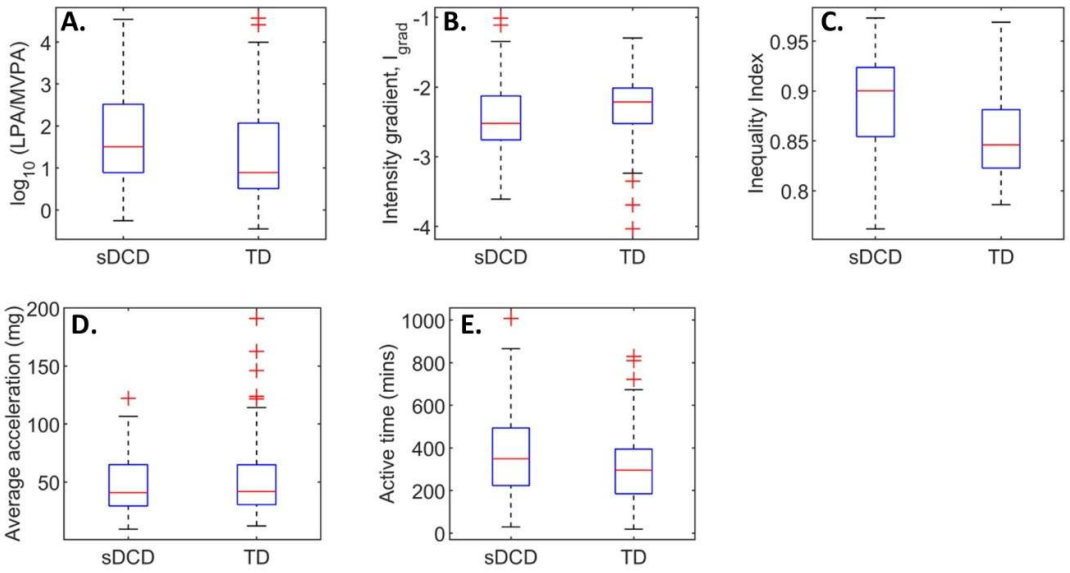

Figure 6 – Boxplots of (A) $\log_{10}$ ratio of minutes spent in LPA and MVPA for sDCD and TD children ($N_{24\ hour\ trace} = 100$, $P = .0046$). (B) Intensity gradient, I_{grad} for sDCD and TD children ($N_{24\ hour} = 100$, $P = .0012$). (C) Inequality index, $I_{\neq}$ for sDCD and TD children ($N_{24\ hour} = 100$, $P = .000033$). (D) Average acceleration for sDCD and TD children ($N_{24\ hour} = 100$, $P = .383$). (E) Total active minutes for sDCD and TD children ($N_{24\ hour} = 100$, $P = .0014$). LPA indicates low-intensity physical activity; MVPA, moderate to vigorous physical activity; sDCD, suspected developmental coordination disorder; TD, typically developing.

The variation in statistical significance for varying sample size is shown in figure 7A. The $I_{\neq}$ index provides high confidence in rejecting the null hypothesis ($p < 0.05$) down to samples sizes ~ 30. In general, the sample size required for a specific level of performance (p-value) is halved for the $I_{\neq}$ index relative to the log ratio or intensity gradient test metrics. The recruitment of participants and acquisition of extended data traces covering 24 hours are particular challenges when working with young children,[44] especially those who may have movement coordination problems.[45] A critical aspect of any test metric for application in this area of study is therefore the cohort size required to achieve reasonable statistical predictive power. The performance of $I_{\neq}$ in this regard is shown in figure 7B, together with results using the log ratio or intensity gradient metrics, or just the total minutes spent in LPA. The inequality index provides the best performance, with > 80% statistical power ($\alpha = 0.05$) possible for cohort numbers > 25.

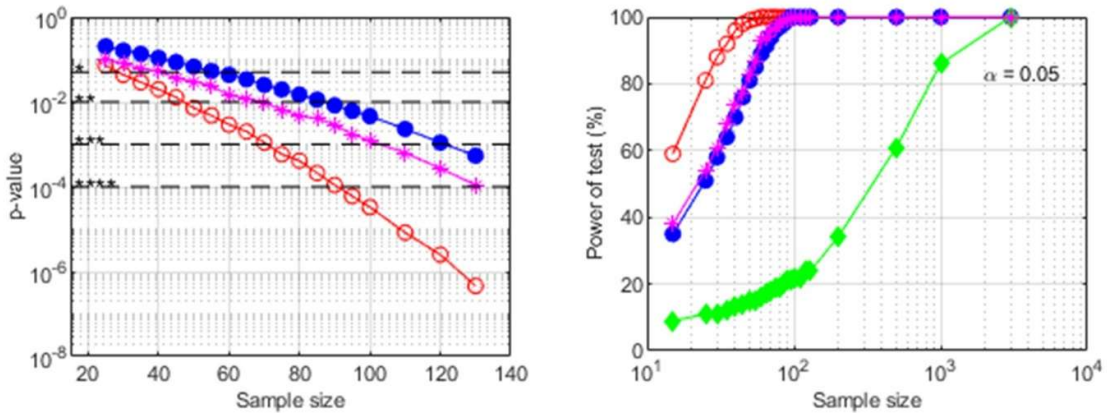

Figure 7 – Test of the null hypothesis of no difference between sDCD and TD children. (A) Probability values as a function of cohort size for the filled circles – $log_{10}\left(\frac{LPA}{MVPA}\right)$, stars – I_{grad}, and open circles – $I_{\neq}$. (B) Statistical power as a function of sample size, for α = .05. filled circles – $log_{10}\left(\frac{LPA}{MVPA}\right)$, stars – I_{grad}, open circles – $I_{\neq}$, and diamonds– LPA minutes. LPA indicates low-intensity physical activity; MVPA, moderate to vigorous physical activity; sDCD, suspected developmental coordination disorder; TD, typically developing.

Conclusions

The rapid enhancement in the capability of sensors provides a technological push to the adoption of more nuanced descriptors of PA. Consider a typical device recording acceleration data over a 24-hour period at an acquisition rate of 100 Hz. This provides > 8 million activity ENMO values, measured with 12-bit resolution. The reduction of this to just 4 measurements of minutes spent in sleep, sedentary behaviour, light physical activity and moderate-to-vigorous activity, constitutes a colossal loss of informative data. More comprehensive statistical indicators are called for, to provide detailed analysis of human movement that can augment the more immediate measures of time spent in activity level bands. To this end, we have presented $I_{\neq}$, an inequality index for PA intensity, which may be viewed as a 'Gini-coefficient for movement'. This captures the full complexity of activity during a 24-hour period through quantification of the time-count variance between bins, in the intensity distribution for the day. At a fundamental level, a frequency distribution may be viewed as an information set, the entropy measure is calculated from the bin-counts of the distribution and quantifies the information content of the data.[46,47] In like manner, the $I_{\neq}$ index captures the information latent in the distribution of activity intensity data. However, when assessing PA our information is limited by the high asymmetry of the intensity distribution – human activity tends to be predominantly sedentary or of low movement intensity.[48-50] The majority of time is recorded against a limited number of low-intensity bins and thus a major part of the distribution carries no information content. Data pre-processing, in the form of a log transformation of the intensity distribution, is therefore required. This enhances information extraction by distributing the 24-hour time allocation across more bins, so providing finer resolution of the data.

In application to children's studies, where high variance between subjects is universal and large sample numbers difficult to collect, $I_{\neq}$ can be transformational in providing reliable statistics for this

sparse and noisy data. Appropriate estimations of statistical power have been critically reviewed in sport and exercise sciences[51] with a call for the use of more stringent approaches. In field type research typical variances are much larger, particularly when focussing on data generated from devices measuring acceleration 24/7. Most research in this area has investigated sample sizes required to validate existing measures of PA, little has been published seeking deeper insight into improving signal processing to help further develop device-based metrology. The application of $I_{\neq}$ to analysis of PA intensity of children across 24-hour periods demonstrates, that for this data, enhanced statistical discrimination is achieved, relative to commonly used metrics such as log ratios and the intensity gradient. This leads to greater statistical power and the key advantage of the $I_{\neq}$ index – it requires smaller cohort numbers to achieve statistical confidence. This enhanced performance stems from the fact that $I_{\neq}$ is based on individual time entries at different activity levels and the wider variation in these across the whole activity range. Local variance is captured by the comparison to a uniform time allocation at each ENMO level, whilst the overall global shape of the distribution is parameterised through the cumulative aspect of the $I_{\neq}$ metric. Whilst the specific probability values obtained from hypothesis testing of children with suspected coordination disorders using $I_{\neq}$ are subject to the limitations of these studies and cannot be assumed to hold for more general application. Nevertheless, the relative performance of $I_{\neq}$ indicates that the inequality index does have the potential to substantially reduce the cohort size required to achieve statistical confidence, compared to alternative test metrics. It is also widely applicable in that its foundation is quite general – a distribution of time across different intensities of PA. It can be implemented with different intensity metrics (eg. ENMO, MAD, MIMS) and makes no assumption on the form of the time distribution.

Alongside statistical testing, PA metrics are also important in providing understanding of the specifics of any changes in activity, i.e. amount of time in movement and levels of exertion. In this regard the PA changes associated with a $\Delta I_{\neq}$ are clearly understandable in a general sense – they quantify greater concentration or more equal distribution of time across activity intensity levels. But a limitation of the

index is that specific values are not immediately interpretable, furthermore subtly different intensity distributions may have the same $I_{\neq}$ value. Further work is therefore needed to discover how changing patterns in the allocation of active time to different intensities of activity affect the numerical value of $I_{\neq}$; and to gain a deeper understanding of the relation of the inequality index to PA volume. This will help to bridge the gap between well-established, immediately interpreted metrics such as total active time or average acceleration and more sophisticated metrics such as $I_{\neq}$ from which it is difficult to gain physical insight.

Acknowledgments

We acknowledge the staff in the Community Paediatric Physiotherapy and Occupational Therapy Department at Singleton Hospital, Swansea, the participants of the Moves-UP project, the schools, and their staff for supporting the project. **Funding source:** This work was supported by funding from the Waterloo Foundation.

References

1. Migueles JH, Aadland E, Andersen LB, Brønd JC, Chastin SF, Hansen BH, Konstabel K, Kvalheim OM, McGregor DE, Rowlands AV, Sabia S, van Hees VT, Walmsley R, Ortega FB. 2022. GRANADA consensus on analytical approaches to assess associations with accelerometer-determined physical behaviours (physical activity, sedentary behaviour and sleep) in epidemiological studies. *Br J Sports Med*. 2022; 56:376-384.

2. Ekelund U, Tarp J, Steene-Johannessen J, Hansen BH, Jefferis B, Fagerland MW, Whincup P, Diaz KM, Hooker SP, Chernofsky A, Larson MG, Spartano N, Vasan RS, Dohrn I-M, Hagströmer M, Edwardson C, Yates T, Shiroma E, Anderssen SA, Lee I-M. Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. *BMJ*. 2019; 366:14570.

3. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nyström C, Mora-Gonzalez J, Löf M, Labayen I, Ruiz JR, Ortega FB. 2017. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports Med*. 2017; 47:1821-1845.

4. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. *Br J Sports Med*. 2014; 48:1019-1023.

5. Bouten CV, Westerterp KR, Verduin M, Janssen JD. Assessment of energy expenditure for physical activity using a triaxial accelerometer. *Med Sci Sports Exerc*. 1994; 26:1516-1523.

6. van Hees VT, Gorzelniak L, Dean-León EC, Eder M, Pias M, Taherian S, Ekelund U, Renström F, Franks PW, Horsch A, Brage S. Separating movement and gravity components in an acceleration

signal and implications for the assessment of human daily physical activity. *PLoS One*. 2013; 8:1–10.

7. Montoye HJ, Washburn R, Servais S, Ertl A, Webster JG. Nagle FJ. Estimation of energy expenditure by a portable accelerometer. *Med Sci Sports Exerc*. 1983; 15:403-407.

8. Plasqui G, Bonomi AG. Westerterp KR. Daily physical activity assessment with accelerometers: new insights and validation studies. *Obes Rev*. 2013: 14:451-462.

9. Hart TL, Ainsworth BE, Tudor-Locke C. Objective and subjective measures of sedentary behavior and physical activity. *Med Sci Sports Exerc*. 2011; 43:449-456.

10. Grgic J, Dumuid D, Bengoechea EG, Shrestha N, Bauman A, Olds T, Pedisic Z. Health outcomes associated with reallocations of time between sleep, sedentary behaviour, and physical activity: a systematic scoping review of isotemporal substitution studies. *Int J Behav Nutr Phys Act*. 2018; 15:1-68.

11. Strain T, Wijndaele K, Dempsey PC, Sharp SJ, Pearce M, Jeon J, Lindsay T, Wareham N, Brage S. Wearable-device-measured physical activity and future health risk. *Nat Med*. 2020; 26:1385-1391.

12. Ekelund U, Steene-Johannessen J, Brown WJ, Fagerland MW, Owen N, Powell KE, Bauman A, Lee IM. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. *Lancet*. 2016; 388:1302-1310.

13. Aadland E, Kvalheim OM, Anderssen SA, Resaland GK, Andersen LB. The multivariate physical activity signature associated with metabolic health in children. *Int J Behav Nutr Phys Act*. 2018; 15:77.

14. Schwendinger F, Infanger D, Lichtenstein E, Hinrichs T, Knaier R, Rowlands AV, Schmidt-Trucksäss A. Intensity or volume: the role of physical activity in longevity. *Eur J Prev Cardiol*. 2025; 32:10-19.

15. Bassett DR, Ainsworth BE, Swartz AM, Strath SJ, O'Brien WL, King GA. Validity of four motion sensors in measuring moderate intensity physical activity. *Med Sci Sports Exerc*. 2000; 32:S471-S480.

16. Fairclough SJ, Taylor S, Rowlands AV, Boddy LM, Noonan RJ. Average acceleration and intensity gradient of primary school children and associations with indicators of health and well-being. *J Sports Sci*. 2019; 37:2159-2167.

17. Rowlands AV, Edwardson CL, Davies MJ, Khunti K, Harrington DM, Yates TOM. Beyond cut points: accelerometer metrics that capture the physical activity profile. *Med Sci Sports Exercise*. 2018; 50:1323-1332.

18. Rowlands AV, Fairclough SJ, Yates T, Edwardson CL, Davies M, Munir F, Khunti K, Stiles VH. Activity intensity, volume, and norms: utility and interpretation of accelerometer metrics. *Med Sci Sports Exerc*. 2019; 51:2410-2422.

19. Lorenz MO, Methods of measuring the concentration of wealth. *Pubs Am Stat Assoc*. 1905; 9:209-219.

20. Atkinson AB. On the measurement of inequality. *J Econ Theory*. 1970; 2:244-263.

21. De Maio FG. Income inequality measures. *J Epidemiol Community Health*. 2007; 61:849-852.

22. Theil H. *Economics and Information Theory*. Amsterdam: North Holland; 1967.

23. Dalton H. The measurement of the inequality of incomes. *Econ J*. 1920; 30:348-361.

24. Aitchison J. The statistical analysis of compositional data. *J R Stat Soc Series B Stat Methodol*. 1982; 44:139-160.

25. Pawlowsky-Glahn V, Egozcue JJ. Compositional data and their analysis: an introduction. *Geol Soc Spec Publ*. 2006; 264:1-10.

26. Rothschild M, Stiglitz JE. Increasing risk: I. A definition. *J Econ Theory*. 1970; 2:225-243.

27. Chastin SFM, Palarea-Albaladejo J, Dontje ML, Skelton DA. Combined effects of time spent in physical activity, sedentary behaviors and sleep on obesity and cardio-metabolic health markers: A novel compositional data analysis approach. *PLos One*. 2015; 10:e0139984.

28. Gini C. Measurement of inequality of incomes. *Econ J*. 1921; 31:124-125.

29. Giorgi GM, Gigliarano C. The Gini concentration index: a review of the inference literature. *J Econ Surv.* 2017; 31:1130-1148.

30. Althoff T, Sosič R, Hicks JL, King AC, Delp SL, Leskovec J. Large-scale physical activity data reveal worldwide activity inequality. *Nature*. 2017; 547:336-9.

31. Chaput JP, Barnes JD, Tremblay MS, Fogelholm M, Hu G, Lambert EV, Maher C, Maia J, Olds T, Onywera V, Sarmiento OL. Inequality in physical activity, sedentary behaviour, sleep duration and risk of obesity in children: a 12-country study. *Obesity science & practice*. 2018; 4:229-37.

32. Swindell N, Starbuck C, Jin S, Barker H, Chesterfield-Thomas GL, Rueda-Hernandez J, Crosby C, Barnes C, Summers HD, Stratton G. The 24-hour movement behaviour compositions of children with and without impaired motor coordination: The Moves-UP project. PLoS One. 2025;20(2):e0319094..

33. den Uil, AR, Janssen M, Busch V, Kat IT, Scholte RH. The relationships between children's motor competence, physical activity, perceived motor competence, physical fitness and weight status in relation to age. *PLoS One*. 2023; 18:e0278438.

34. Stodden DF, Goodway JD, Langendorfer SJ, Roberton MA, Rudisill ME, Garcia C, Garcia LE. A developmental perspective on the role of motor skill competence in physical activity: An emergent relationship. *Quest*. 2008; 60:290-306.

35. Tamplain P, Cairney J. Low motor competence or developmental coordination disorder? An overview and framework to understand motor difficulties in children. *Curr Dev Disord Rep*. 2024; 11:1-7.

36. Fairclough SJ, Rowlands A V., del Pozo Cruz B, Crotti M, Foweather L, Graves LEF, Hurter L, Jones O, MacDonald M, McCann DA, Miller C, Noonan RJ, Owen MB, Rudd JR, Taylor SL, Tyler R, Boddy LM. Reference values for wrist-worn accelerometer physical activity metrics in England children and adolescents. *Int J Behav Nutr Phys Act*. 2023; 20:35.

37. Hildebrand M, Van Hees VT, Hansen BH, Ekelund U. Age group comparability of raw accelerometer output from wrist-and hip-worn monitors. *Med Sci Sports Exerc*. 2014; 46:1816–24.

38. Hurter L, Fairclough SJ, Knowles ZR, Porcellato LA, Cooper-Ryan AM, Boddy, LM. Establishing Raw Acceleration Thresholds to Classify Sedentary and Stationary Behaviour in Children. *Children*. 2018; 5: 172.

39. Wilson BN, Crawford SG, Green D, Roberts G, Aylott A, Kaplan BJ. Psychometric properties of the revised developmental coordination disorder questionnaire. *Phys Occup Ther Pediatr*. 2009; 29:182–202.

40. Wilk MB, Gnanadesikan R. Probability plotting methods for the analysis for the analysis of data. *Biometrika*, 1968; 55:1–17.

41. Cohen J. *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge; 2013.

42. Mumby PJ. Statistical power of non-parametric tests: A quick guide for designing sampling strategies. *Mar Pollut Bull*. 2002; 44:85-87.

43. Cox NJ. Speaking stata: Logarithmic binning and labelling. *Stata J.* 2018; 18:262-286.

44. Welk GJ, Corbin CB, Dale D. Measurement issues in the assessment of physical activity in children. *Res Q Exerc Sport*. 2000; 71:59-73.

45. Visser J, Geuze RH, Kalverboer AF. The relationship between physical growth, the level of activity and the development of motor skills in adolescence: Differences between children with DCD and controls. *Hum Mov Sci*. 1998; 17:573-608.

46. Shorrocks AF. The class of additively decomposable inequality measures. *Econometrica*. 1980; 48:613-625.

47. Pielou EC. The measurement of diversity in different types of biological collections. *J Theor Biol*. 1966; 13:131-144.

48. Owen N, Healy GN, Matthews CE, Dunstan DW. Too much sitting: the population health science of sedentary behavior. *Exerc Sport Sci Rev*. 2010; 38:105-113.

49. Hayes M, Chustek M, Heshka S, Wang Z, Pietrobelli A, Heymsfield SB. Low physical activity levels of modern Homo sapiens among free-ranging mammals. *Int J Obes*. 2005; 29:151-156.

50. Rhodes RE, Mark RS, Temmel CP. Adult sedentary behavior: a systematic review. *Am J Prev Med*. 2012; 42:e3-e28.

51. Abt G, Boreham C, Davison G, Jackson R, Nevill A, Wallace E, Williams M. Power, precision, and sample size estimation in sport and exercise science research. *J Sports Sci*. 2020; 38:1933-5.