Annals of Clinical and Translational Neurology

RESEARCH ARTICLE OPEN ACCESS

Assessing a Mitochondrial Disease Treatment via a Novel Statistical Technique for Accelerometer Data

Ian W. McKeague[1] | Kristin Engelstad[2] | Yue Ge[1] | Amber Tucker[2] | Shufang Li[2] | Jasim Uddin[2] | Ziwei Zhao[3] | John L. P. Thompson[1] | Michio Hirano[2]

[1]Department of Biostatistics, Columbia University, New York, New York, USA | [2]Department of Neurology, Columbia University, New York, New York, USA | [3]Centre for Quantitative Medicine, Duke-NUS Medical School, Singapore, Singapore

Correspondence: Michio Hirano (mh29@columbia.edu)

Received: 6 May 2025 | **Revised:** 2 August 2025 | **Accepted:** 5 August 2025

Funding: This work was supported by National Institutes of Health (1R01 AG062401 and U54 NS078059).

Keywords: accelerometer | mitochondrial disease | outcome measure | thymidine kinase 2 deficiency

ABSTRACT

Objective: Therapeutic development for mitochondrial diseases, rare genetic disorders with pathogenic defects of oxidative phosphorylation, is hindered by unsatisfactory outcome measures. To address this problem, we provide the first clinical application of a novel, bias-adjusted outcome measure of acceleration across a range of subjects' activities to assess nucleoside therapy for thymidine kinase 2 deficiency, an ultra-rare autosomal recessive mitochondrial disease.

Methods: Data were collected from treated patients in an ongoing phase 2 clinical trial who served as their own controls. If there is a treatment effect, time-in-activity curves for these patients will increase over successive clinic visits. We used a combination of functional data analysis and longitudinal mixed-effects linear regression, adjusted for age and gender, to test for the effect of treatment length on time-in-activity.

Results: For 14 patients with at least two assessments 6 months apart, we found a significant overall improvement of time-in-activity due to treatment. Improvement was especially significant at two individual activity levels within the range (0.14 and 2 g). In longitudinal analyses, using data on time-in-activity at these two levels for all clinic visits of 19 subjects, the effect of treatment length on time-in-activity was highly significant at both 0.14 g (0.04, CI 0.01–0.08, $p = 0.023$) and 2 g (0.01, 0.00–0.02, $p = 0.013$).

Interpretation: This small-N exploratory analysis using a new accelerometer-based activity measure featuring powerful data reduction and adjustment for circadian rhythms and other biases finds that nucleoside therapy may increase activity levels in thymidine kinase 2 deficiency patients.

1 | Introduction

Mitochondrial diseases are genetic disorders due to oxidative phosphorylation defects resulting from primary mutations in mitochondrial DNA (mtDNA) or nuclear DNA [1]. Onset of thymidine kinase 2 deficiency (TK2d) can occur at any age. Patients with initial TK2d manifestations in infancy present a severe, rapidly progressive phenotype, whereas those with late onset greater than 12 years old demonstrate variably slower progression [2, 3]. Onset has been classified into 3 groups, based on the presentation of myopathy. Onset in infancy (≤1 years) has a median post-onset survival of less than 1 year, whereas childhood-onset (>1 to ≤12 years old) and late onset (>12 years old) have a median survival of at least 12 years [4]. Although progressive myopathy with progressive limb weakness, dysphagia, and restrictive lung disease is the predominant clinical feature of TK2d, the disorder can have multisystemic manifestations, with seizures, cognitive impairment, hearing loss, peripheral neuropathy, and cardiopathy [4–6]. Due to the myopathic weakness and other manifestations, TK2d is often debilitating, with

progressive loss of motor functions that can impair the ability to perform routine activities of daily living.

TK2 is a mitochondrial matrix protein that is a key component of the mitochondrial deoxynucleotide salvage pathway via phosphorylation of [deoxy]thymidine (dT) and deoxycytidine (dC) pyrimidine nucleosides, generating thymidine and deoxycytidine monophosphate (dTMP + dCMP), which are subsequently converted to deoxynucleoside triphosphates (dNTPs) for the synthesis of mtDNA [2]. While there is currently no FDA approved treatment for TK2d, nucleoside therapy (NT) is a promising treatment being investigated through a clinical trial (TK0102; NCT03845712) to determine its safety and efficacy [7].

Due to the etiological heterogeneity and clinical diversity of mitochondrial diseases, a variety of primary clinical outcomes have been utilized in clinical trials including: composite scores of manifestations (e.g., Newcastle Mitochondrial Disease Adult Scale), seizure frequency, fatigue, muscle functions (e.g., 6 min walk test, cycle ergometry), and quality of life measures [8–10]. Assessment of accelerometer readings may be an additional useful outcome measure to include in the analysis of potential mitochondrial disease treatments designed to improve motor function. The most prominent clinical manifestation of TK2d is muscle weakness, which can be sensed and monitored by accelerometers. Accelerometers are wearable devices that can provide massive, objective, and near-continuous measurements in response to gravity and physical activities. Such real-time human activity data is valuable in monitoring long-term physical status and assessing biomechanical parameters in healthcare. However, analyzing human activity data from accelerometers is often challenging due to the large number of observations in time-series data, the temporal property of the observations, and the undefined relationship between observed data and actual human movements [11]. Recently, Chang and McKeague [12] proposed a nonparametric inference framework to compare accelerometer data between groups of subjects. We apply this approach to TK2d patients in a Phase 2 clinical trial that is assessing the efficacy of NT using survival and motor milestones as outcome measures. Our hypothesis is that treatment with NT may improve daily motor functions and result in higher activity levels over time.

2 | Subjects and Methods

Our study enrolled 14 participants with genetically confirmed TK2d who were already being treated with NT at Columbia University Irving Medical Center (CUIMC) in a phase 2, open-label clinical trial (TK0102: 12 subjects) or expanded access protocols (EAP: 2 subjects). An additional 5 participants were enrolled but were only assessed on one occasion and could not be used for the primary analysis. Informed consent was obtained from all subjects or legal guardians under a CUIMC Institutional Review Board approved protocol (AAAS6805). Subjects 7–17 years old were required to provide signed consent. Each participant was expected to wear the activPAL accelerometer device applied to the inner mid-thigh of the dominant leg for 7 consecutive days at approximately 6-month intervals, which often coincided with the clinical trial or EAP visit schedule. Baseline characteristics of the 14 subjects are listed in Table 1. The second-last column indicates post-baseline visits with accelerometer data of acceptable quality (i.e., with at least three consecutive days of data). The last column indicates the number of days from V1 to these visits. The primary analysis compares the last visit (indicated by "a") to the first (V1). An extended analysis uses data of acceptable quality from *any* subsequent visit after the first visit (i.e., all visits listed in the second-to-last column).

2.1 | Data Processing

The activPal accelerometer is a research-quality miniature electronic logger developed by PAL Technologies to quantify free-living daily activities (PAL Technologies Ltd., 2023) [13] based on acceleration data. The device provides raw (x,y,z)-components of acceleration in units of g (acceleration due to gravity) at a frequency of 20 Hz.

Let the (x,y,z)-components of acceleration observed at time t be denoted $(X(t), Y(t), Z(t))$, where the time index t varies over the monitoring period with 20 values per second. The Euclidean norm $r(t) = \sqrt{X(t)^2 + Y(t)^2 + Z(t)^2}$ was then computed at each time point t, and a unit offset applied to remove the effect of gravity, resulting in the "corrected" accelerometer reading $S(t) = \max\{r(t)\text{-}1,0\}$, which, as a function of t, is referred to in the sequel as *activity data*.

To avoid bias in the activity data, if the device was removed early in follow-up, the monitoring period was clipped (truncated) to a multiple of 24 h (throughout which the device was observed to be continuously in use). A plot of the activity data for a particular subject (Figure 1a) illustrates the problem. Figure 1b shows the truncated version of the activity data. Figure 1c illustrates a further "data cleaning" step which filters out low level activity (attributable to noise in the data) by only recording the maximum activity over each minute. The final step in the data processing was to renormalize the activity data for subjects having less than 7 days of data by rescaling the time axis to represent exactly 7 days.

2.2 | Time-In-Activity Curves

Chang and McKeague [12] introduced a flexible approach to the analysis of wearable device data focusing attention on curves that describe *time-in-activity* ("occupation time curves"), the total amount of time over the monitoring period during which the activity (acceleration) exceeds a given level, as that level varies. The rationale for using time-in-activity curves, rather than the much more complex activity data themselves, is that it is possible thereby to compare the time spent in different levels of activity without having to adjust for differences in their temporal patterns (e.g., circadian rhythm). This approach is applicable in the present setting by results of Chang and McKeague [12].

The *time-in-activity curve* $T(a)$ is the total amount of time that the activity level $S(t)$ exceeds the level a, as a varies over the range of the accelerometer readings (which is in the range 0–2 g

TABLE 1 | Demographic and clinical summary (14 subjects).

ID	Gender	Age (years) at			Dose (mg/kg/d)	Pathogenic variants	Ambulatory status (A or NA)	V/R support	Feeding status	Visits compared to V1	Days from V1[a]
		Symptom onset	Treatment start	Last follow-up							
2051	M	2	9	9	400	c.323C>T (p.T108M) Homozygous	A	No	PO	V3, V4[a]	272, 469[a]
2033	F	1	4	11	400	c.323C>T (p.T108M); c.361C>A, p.H121N)	A	No	PO	V3, V5[a]	250, 612[a]
2041	M	1	1	3	400	c.323C>T (p.T108M); c.644T>C (p.L215P)	A	No	NG tube	V2[a]	287[a]
2042	M	1	1	3	400	c.323C>T (p.T108M); c.644 T>C (p.L215P)	A	No	NG tube	V2, V3[a]	292, 476[a]
2024	F	12	15	18	130	c.173A>G (p.N58S) Homozygous	A	BiPAP at night	G tube and PO	V2, V3[a]	272, 453[a]
2025	M	3	14	15	130	c.173A>G (p.N58S) Homozygous	NA from 12 October 2021 to present due to fracture.	No	PO	V3[a]	453[a]
2039	F	5	17	19	400	c.604_606delAAG (p.K202del); c.575G>A (p.R192K)	A	CPAP at night	PO	V4[a]	550[a]
2049	M	1	2	3	400	c.698 T>C (p.L233P); c. 623A>G (p. Y208C)	NA	~20 h/day	G tube	V2, V3, V6[a]	98, 178, 382[a]
EA2	M	1.5	3	5	400	c.323C>T (p.T108M) Homozygous	NA	Invasive ventilation from age 41 months	J tube	V2[a]	214[a]
EA3	M	0.83 (10 months)	1	1	400	c.547C>T (p.R183W) and c.447_448dup AA	NA	BiPAP-16 h overnight from age 16 month	G tube	V4[a]	477[a]
2043	F	1	2	4	400	c.547C>T (p.R183W) Homozygous	A	No	PO (G tube for IP only)	V2, V3[a]	238, 422[a]
2045	F	1	3	5	400	c.323C>T (p.T108M); c.272G>A (p.G91N)	A	No	PO	V2[a]	161[a]

(Continues)

TABLE 1 | (Continued)

ID	Gender	Age (years) at Symptom onset	Age (years) at Treatment start	Age (years) at Last follow-up	Dose (mg/kg/d)	Pathogenic variants	Ambulatory status (A or NA)	V/R support	Feeding status	Visits compared to V1	Days from V1[a]
2013	M	3	8	11	400	c.361C>A (p.H121N); c.549C>T (p.R183W)	NA	CPAP at night	PO	V2[a]	247[a]
2014	F	2	4	7	400	c.361C>A (p.H121N); c.549C>T (p.R183W)	A	No	PO	V2[a]	247[a]

Note: Dose is milligrams/kg body weight/day (mg/kg/d).
Abbreviations: A, ambulatory; NA, non-ambulatory; V/R, ventillatory/respiratory.
[a]These visits were compared to V1 in the primary analysis.

in the sequel). In practice, the calculation of $T(a)$ is on a fine grid of values of the acceleration a, corresponding to a regular grid of values of say 500 points on the y-axis in Figure 1c. The time-in-activity curves for two subjects after two selected visits are shown in Figure 2.

2.3 | Treatment Effect Curves

We address the problem of assessing treatment efficacy in terms of the *difference* between the time-in-activity curve at a given visit and the time-in-activity curve at an earlier visit. If the treatment has an effect, then we expect the time-in-activity curve to increase with each successive visit, at least over some range of activity levels.

3 | Results

Figure 3 shows the differences between the time-in-activity curves between last (indicated by "a" in Table 1) and first visits for the 14 subjects used in the primary analysis. This "treatment effect" is plotted on the y-axis in units of hours over a week of observation, versus the activity level over the range0 to 2 g on the x-axis. Observe that a high proportion of the treatment effect curves fall above zero at low levels of activity (around $a = 0.14$ g) as well as at high levels of activity (around $a = 2$ g).

3.1 | Testing for the Presence of a Treatment Effect

Figure 3 shows the simultaneous 95% confidence band for the mean treatment effect based on the data plotted in Figure 3. This band was developed specifically for analyzing wearable device data by Chang and McKeague [12] and is implemented in the R package fdEL developed by Chang [14]. Notice that the pointwise estimates of the mean (solid line) are mostly above zero, and the lower limits of the confidence band at activity levels 0.14 and 2 g are both greater than zero. This provides strong evidence of the presence of treatment efficacy at both low and high activity levels. The use of a *simultaneous* confidence band is crucial in this analysis because it is then statistically valid to search across activity levels without losing confidence in the coverage.

Figure 5 shows the simultaneous 95% confidence band for the mean treatment effect inversely weighted by the number of years between the first and last visits. Figure 6 shows the simultaneous 95% confidence band for the mean treatment effect based on the extended analysis of data from all 21 post-baseline visits of the 14 subjects listed in Table 1.

In summary, for the 14 patients studied in the primary analysis, the test based on the 95% confidence band in Figure 3 shows a significant overall improvement in the time-in activity curves. Improvement was especially significant at two specific activity levels within the range of 0 to 2 g, namely at 0.14 and 2.00 g. These effects are confirmed in both the inversely weighted confidence band (Figure 6) for the same set of 14 subjects and also in the extended analysis (using data on all 21 subsequent visits for the 14 subjects, Figure 6).

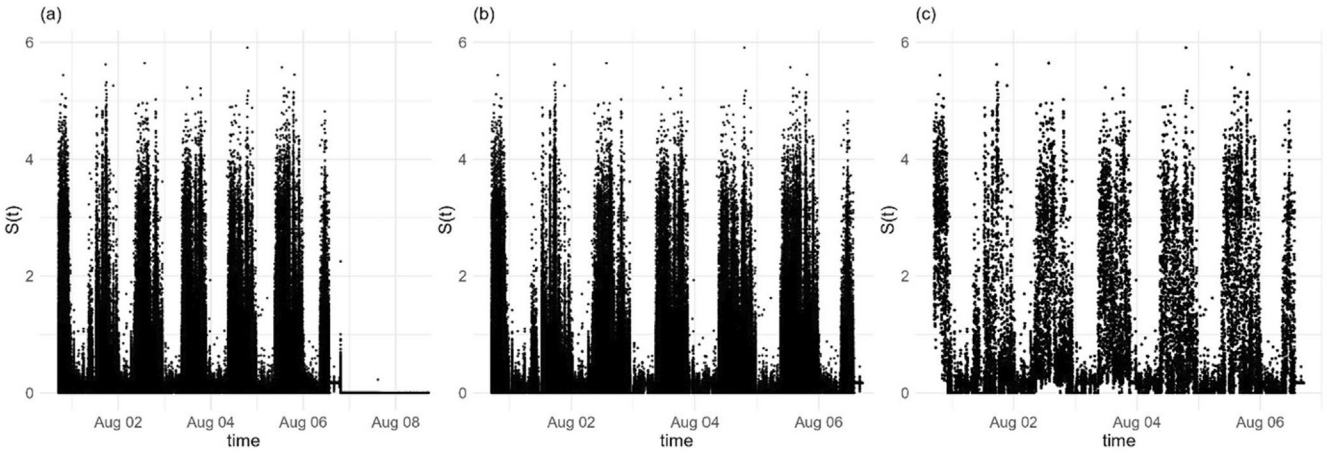

FIGURE 1 | Activity data (acceleration in units of g) plotted as a function of calendar time for subject 2051, monitored after visit 3. (a) Shows a period of inactivity at the end of follow-up, (b) shows the truncated activity data, after the period of inactivity is removed; (c) shows the filtered activity data after maximizing the acceleration over each minute.

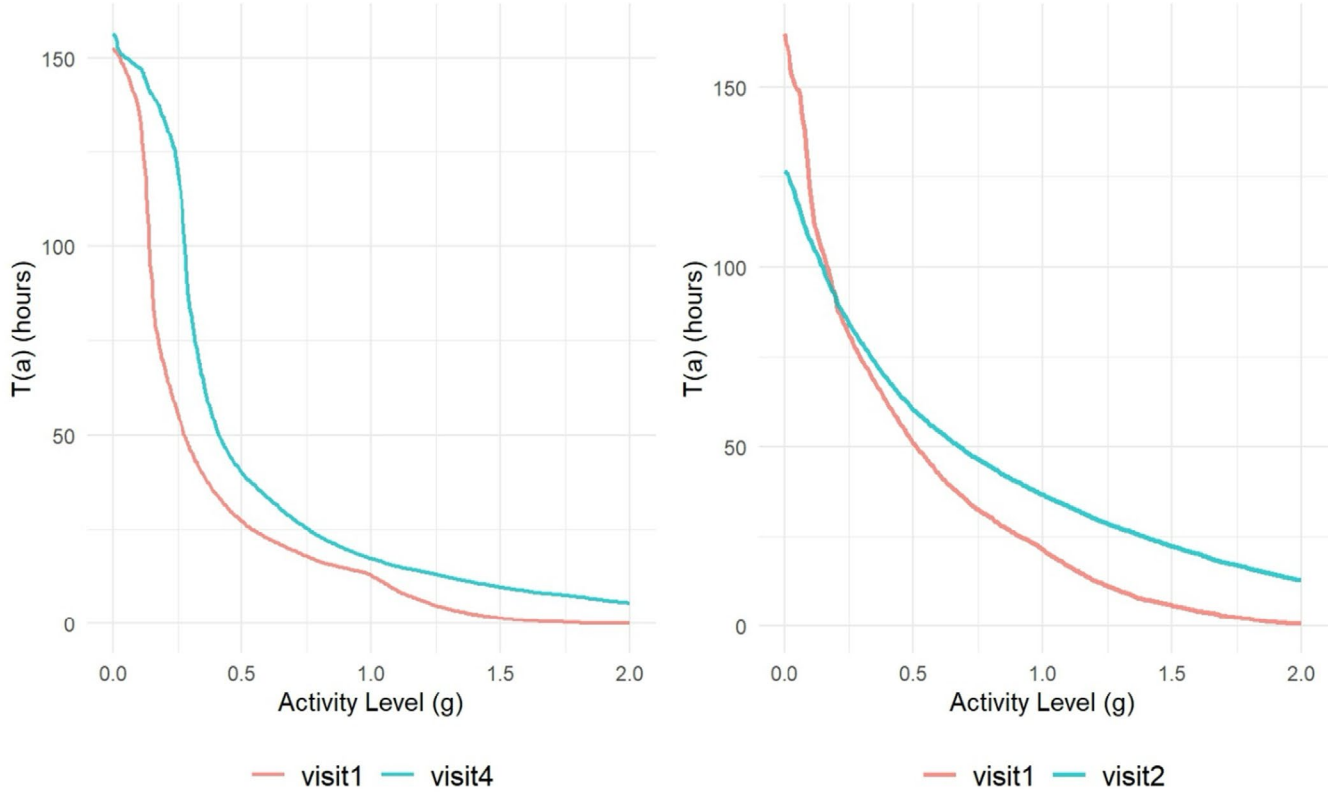

FIGURE 2 | Time-in-activity curves $T(a)$ (in hours over a week) color-coded by visit number and plotted against the activity level (a) on a grid of 500 points in 0–2 g, for subject 2039 (left panel) and subject 2041 (right panel).

3.2 | Longitudinal Analysis for Low and High Activity Levels

Having identified the activity levels $a = 0.14$ g and $a = 2$ g to be of interest, we proceed to develop a longitudinal model that can assess the treatment effect (on time-in-activity) over a sequence of repeated visits (rather than just two visits). This is done by fitting separate random intercept linear regression models for the time-in-activity values $T(0.14)$ and $T(2.00)$, with a random intercept used to take into account within-subject correlation over different visits of the same subject. The model includes fixed effects for treatment length in days (which gives the slope parameter of principal interest), age when treated, and gender [1 for male, 0 for female], as well as the (Gaussian) random intercept for each subject.

The analysis reported in Table 2 is based on using the available time-in-activity curves for all visits of 19 subjects (the 14 original subjects and the 5 additional subjects who have only a single

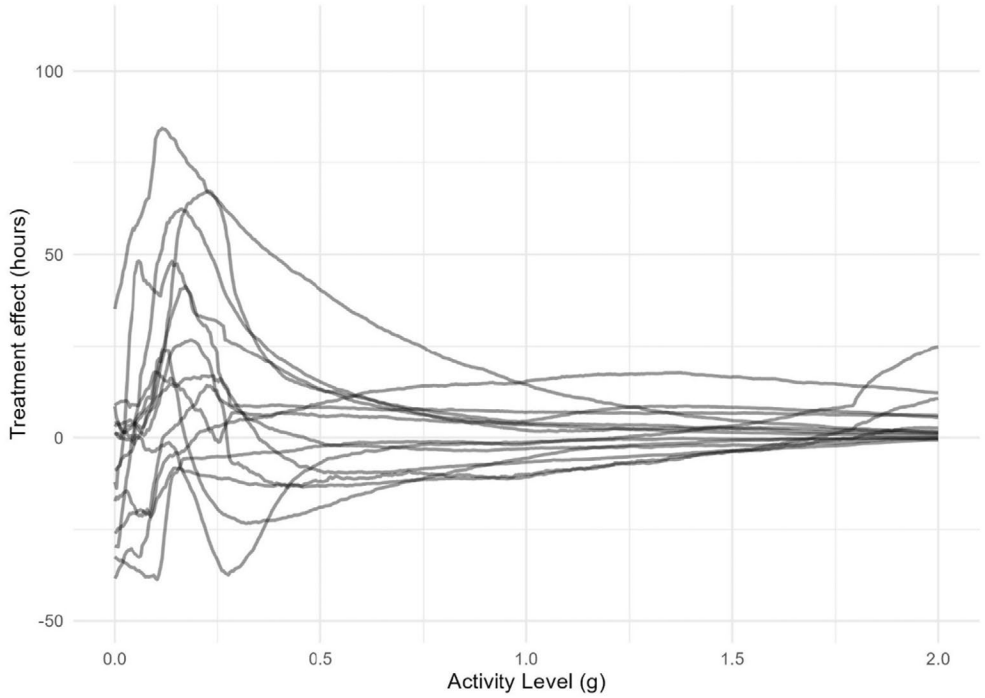

FIGURE 3 | Primary analysis data. Treatment effect curves (differences in time-in-activity curves between first and last visits), plotted as a function of activity level (in units of *g*), for 14 subjects.

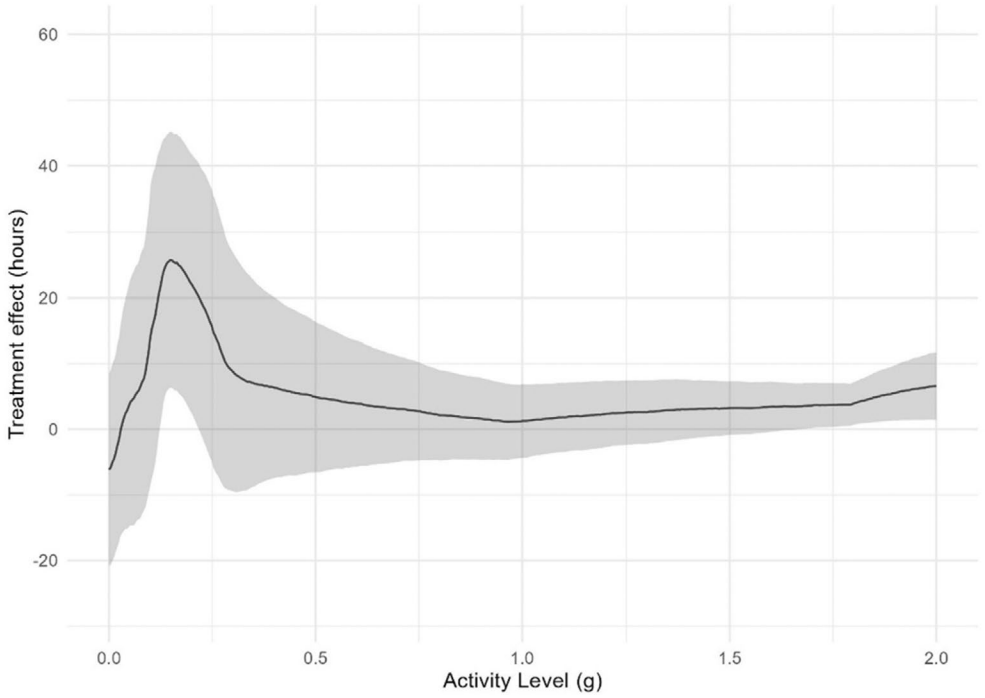

FIGURE 4 | Primary analysis. 95% simultaneous confidence band for mean treatment effect (hours over a week) plotted as a function of activity level, for the 14 subjects with treatment effects plotted in Figure 4.

visit). The results indicate highly significant effects of treatment length on time-in-activity at both lower and higher activity levels. The point estimate of the slope parameter for the effect of treatment length (in days after V1) on $T(0.14)$ is 0.04 (CI = 0.01–0.08, $p = 0.023$). For $T(2.00)$, the point estimate is 0.01 (CI = 0.00–0.02, $p = 0.013$). As mentioned above, both regression analyses were adjusted for age and gender as fixed effects; although neither age nor gender was found to be significant, they were included in the analysis as they may well be important as data on more subjects become available.

4 | Discussion

Mitochondrial diseases are rare genetic disorders with pathogenic defects of oxidative phosphorylation. TK2 deficiency, an ultra-rare autosomal recessive mitochondrial disease, manifests predominantly as a myopathy impairing motor functions including ambulation, breathing, and swallowing. In an exploratory analysis, we examined the efficacy of nucleoside therapy for 14 TK2-deficient patients over a number of clinic visits in an ongoing Phase 2 clinical trial [7] or expanded access protocols at a single site, using accelerometer data collected with activPal, a research-quality miniature electronic logger developed by PAL

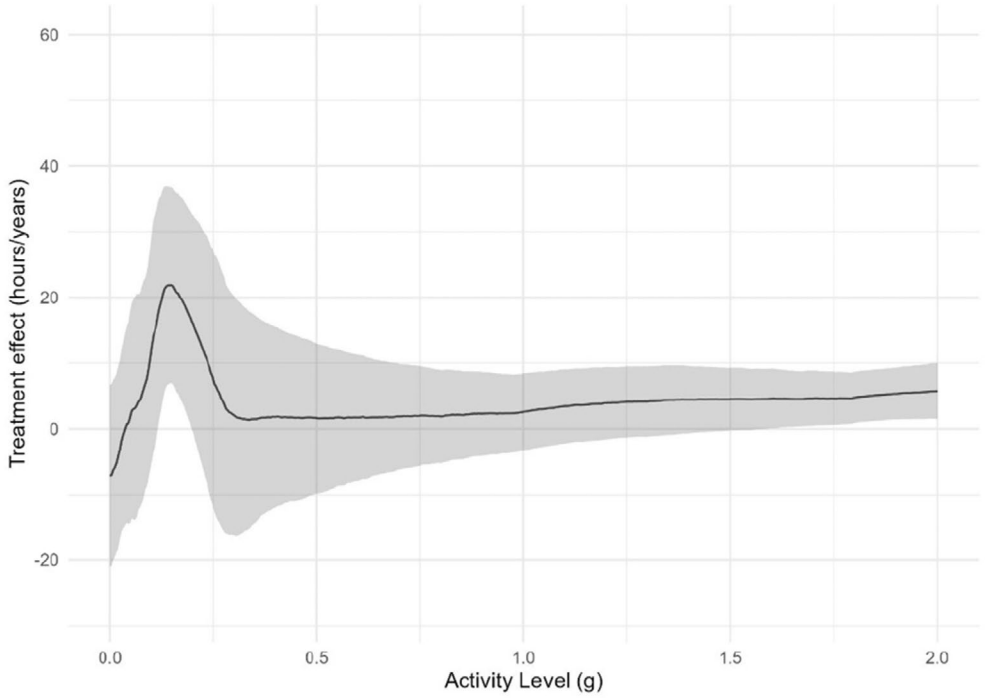

FIGURE 5 | Inversely weighted analysis. 95% simultaneous confidence band for the inversely-weighted mean treatment effect (in hours over a week, inversely weighted by years between visits being compared) plotted as a function of activity level, for the 14 subjects with treatment effects plotted in Figure 4.

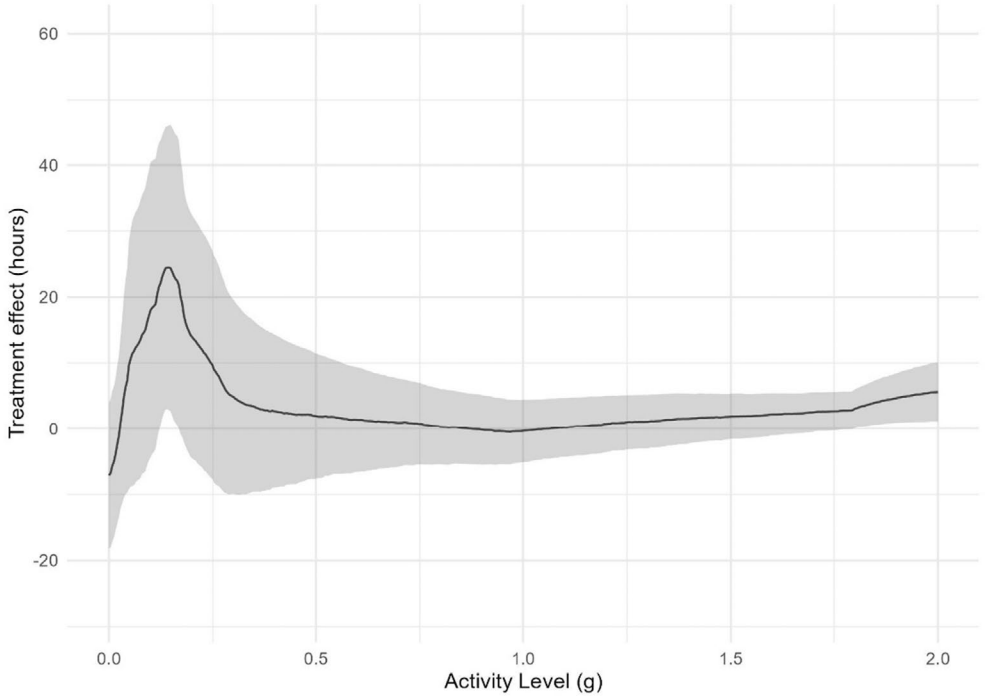

FIGURE 6 | Extended analysis. 95% simultaneous confidence band for the mean treatment effect (in hours over a week) plotted as a function of activity level, based on the data for all subsequent visits of 14 subjects.

TABLE 2 | Longitudinal analysis for low and high activity levels, showing a highly significant effect of treatment length (in days) on time-in-activity, for 19 subjects (40 observations).

	Activity level = 0.14 g			Activity level = 2 g		
Predictors	**Estimates**	**CI**	***p***	**Estimates**	**CI**	***p***
Intercept	93.81	77.65 to 109.97	< 0.001	16.15	6.86 to 25.43	0.001
Gender	−6.82	−22.28 to 8.64	0.376	−8.50	−18.58 to 1.58	0.096
Age	0.48	−0.78 to 1.73	0.444	−0.55	−1.29 to 0.19	0.140
Treatment length	0.04	0.01 to 0.08	0.023	0.01	0.00 to 0.02	0.013

Technologies. We hypothesized that activity levels would increase over time in patients under the nucleoside therapy. Using an innovative nonparametric inference framework for functional data analysis [12], and utilizing patients as their own controls, in our primary analysis we find a 95% confidence band for the mean difference in activity patterns over the week following the first lab visit and subsequent visit. The confidence level holds simultaneously over levels of acceleration in the range 0 to 2 g. The results indicate the presence of a significant increase in activity at low (0.14 g) and high (2 g) levels of acceleration. Secondary analyses incorporating additional visits support this conclusion.

This small-N exploratory analysis using a new acceleration measure featuring powerful data reduction and adjustment for circadian rhythms and other biases finds that NT therapy may improve activity levels in TK2 patients and is consistent with the reported restoration of motor milestones [15]. Given the small patient numbers in ultra-rare diseases and the poor performance of existing outcomes, this finding is striking and potentially a landmark. We are seeking the necessary confirmation and comparisons with other outcomes in further work with NT-treated TK2 patients. Our outcome and functional data analysis methodology are also widely applicable in many other areas, such as natural history studies and intervention trials for other MDs, and development of a possible primary outcome biomarker in clinical trials for MDs and other diseases, prevalent as well as rare. The potential for this new approach is substantial.

Author Contributions

I.M. designed the inferential method to estimate treatment effects based on accelerometer (activPal) data, developed the statistical analysis, interpreted the statistical findings, and took the lead in writing the manuscript. K.E. drafted and implemented the clinical protocol and managed the dataset of raw activPal measurements. Z.Z. and Y.G. contributed to writing and running the R code that produced the statistical results. A.T., S.L., and J.U. collected and curated the activPal data. S.L. also assessed patients clinically, and A.T. contributed to the draft of the clinical text. J.L.P.T. contributed to the design and interpretation of the analyses, and critically revised the manuscript. M.H. was site PI of the TK0102 clinical trial (NCT03845712), designed the clinical analyses, and revised the manuscript critically. He identified subjects, supervised clinical activities, and supervised activPal administration and IRB protocol development.

Acknowledgements

The authors greatly appreciated the participation of the subjects and their families. The work of Ian McKeague and Michio Hirano was partially supported by NIH (Grants 1R01 AG062401 and U54 NS078059). The authors thank Jeff Goldsmith and Chenxi Liu for help during the early stages of the project.

Conflicts of Interest

Dr. Michio Hirano has been a paid consultant to Modis Therapeutics Inc. (a subsidiary of Zogenix/UCB). He has received honoraria for serving on Modis Therapeutics Advisory Boards. This relationship is *de minimus* for Columbia University Irving Medical Center. Dr. Hirano is also co-inventor in Columbia University's patent "Deoxynucleoside Therapy for Diseases Caused by Unbalanced Nucleotide Pools including Mitochondrial DNA Depletion Syndromes" (US Patent 10,471,087), which has been licensed to Modis Therapeutics and has received royalty payment from Modis Therapeutics; this relationship is monitored by an unconflicted external academic researcher. All other authors report no Conflicts of Interest related to this work.

Data Availability Statement

Data are available upon request to the corresponding author.

References

1. G. S. Gorman, P. F. Chinnery, S. DiMauro, et al., "Mitochondrial Diseases," *Nature Reviews Disease Primers* 2 (2016): 16080.

2. A. Berardo, C. Dominguez-Gonzalez, K. Engelstad, and M. Hirano, "Advances in Thymidine Kinase 2 Deficiency: Clinical Aspects, Translational Progress, and Emerging Therapies," *Journal of Neuromuscular Diseases* 9 (2022): 225–235.

3. A. Saada, A. Shaag, H. Mandel, Y. Nevo, S. Eriksson, and O. Elpeleg, "Mutant Mitochondrial Thymidine Kinase in Mitochondrial DNA Depletion Myopathy," *Nature Genetics* 29 (2001): 342–344.

4. C. Garone, R. W. Taylor, A. Nascimento, et al., "Retrospective Natural History of Thymidine Kinase 2 Deficiency," *Journal of Medical Genetics* 55 (2018): 515–521.

5. J. Wang, E. Kim, H. Dai, et al., "Clinical and Molecular Spectrum of Thymidine Kinase 2-Related mtDNA Maintenance Defect," *Molecular Genetics and Metabolism* 124 (2018): 124–130.

6. C. Dominguez-Gonzalez, A. Hernandez-Voth, L. B. Guerrero, et al., "Metrics of Progression and Prognosis in Untreated Adults With Thymidine Kinase 2 Deficiency: An Observational Study," *Neuromuscular Disorders* 32 (2022): 728–735.

7. M. Hirano, C. Garone, R. Haas, et al., eds., "Survival Analyses in Patients With Thymidine Kinase 2 Deficiency Aged ≤12 Years at Symptom Onset Who Received Pyrimidine Nucleos(t)ide Therapy," in *Muscular Dystrophy Association Clinical & Scientific Conference* (Muscular Dystrophy Association, 2025), https://www.mdaconference.org/abstract-library/survival-and-functional-outcomes-in-patients-with-thymidine-kinase-2-deficiency-aged-12-years-at-symptom-onset-who-received-pyrimidine-nucleostides/.

8. A. Karaa, E. Bertini, V. Carelli, et al., "Efficacy and Safety of Elamipretide in Individuals With Primary Mitochondrial Myopathy: The MMPOWER-3 Randomized Clinical Trial," *Neurology* 101 (2023): e238–e252.

9. K. L. Madsen, A. E. Buch, B. H. Cohen, et al., "Safety and Efficacy of Omaveloxolone in Patients With Mitochondrial Myopathy: MOTOR Trial," *Neurology* 94 (2020): e687–e698.

10. A. Karaa and T. Klopstock, "Clinical Trials in Mitochondrial Diseases," in *Mitochondrial Diseases*, ed. R. Horvath, M. Hirano, and P. F. Chinnery (Elsevier, 2023), 229–250.

11. J. Brownlee, "A Gentle Introduction to a Standard Human Activity Recognition Problem," MachineLearningMastery.com (2019) [March 16, 2023], https://machinelearningmastery.com/how-to-load-and-explore-a-standard-human-activity-recognition-problem/.

12. H.-W. Chang and I. W. McKeague, "Empirical Likelihood-Based Inference for Functional Means With Application to Wearable Device Data," *Journal of the Royal Statistical Society Series B: Statistical Methodology* 84 (2022): 1947–1968.

13. Ltd. PT, "Why Activpal?," (2023), https://www.palt.com/why-activpal/.

14. H.-W. Chang, "fdEL [Electronic Resource: R Code]," (2022), https://github.com/news11/fdEL.

15. C. Garone, M. Hirano, R. Haas, et al., eds., "Functional Outcomes in Patients With Thymidine Kinase 2 Deficiency Aged ≤12 Years at Symptom Onset Who Received Pyrimidine Nucleos(t)ide Therapy," in *Muscular Dystrophy Association Clinical & Scientific Conference* (Muscular Dystrophy Association, 2025), https://www.mdaconference.org/abstract-library/functional-outcomes-in-patients-with-thymidine-kinase-2-deficiency-aged-%e2%89%a412-years-at-symptom-onset-who-received-pyrimidine-nucleostide-therapy/.

16. H.-W. Chang and I. W. McKeague, "Empirical Likelihood in Functional Data Analysis," *Annual Review of Statistics and Its Application* 12 (2025): 425–448 [Open source].

Appendix A

A recent article of Chang and McKeague [16] provides an open-source and non-technical survey of the use of empirical likelihood methods in functional data analysis, the basis for the analytical approach we have utilized, and compares such methods to previously established functional data analysis techniques.

Empirical likelihood is a nonparametric likelihood ratio procedure that has various advantages over traditional (Wald-type) procedures, notably in terms of the accuracy of confidence regions and the fact that it does not require smoothing techniques typically needed in functional data analysis. The key technical condition that justifies our use of empirical likelihood is that the observed functions are of *bounded variation* (allowing for jump discontinuities as well). This condition is satisfied by the treatment effect curves because these functions are differences of *time-in-activity curves*, which are decreasing functions of activity level. This facilitates the construction of simultaneous confidence bands for the mean treatment effect over a range of activity levels, a crucial ingredient of our approach.