Published in final edited form as:

J Sci Med Sport. 2015 November ; 18(6): 667–672. doi:10.1016/j.jsams.2014.09.003.

A multilevel approach to examining time-specific effects in accelerometer-assessed physical activity

Hannah G. Lawman^{a,*}, M. Lee Van Horn^b, Dawn K. Wilson^b, and Russell R. Pate^c

^aCenter for Obesity Research and Education, Temple University, USA

^bDepartment of Psychology, University of South Carolina, USA

^cDepartment of Exercise Science, Arnold School of Public Health, University of South Carolina, USA

Abstract

Objectives—Popular methods for analyzing accelerometer data often use a single physical activity outcome variable such as average-weekly or total physical activity. These approaches limit the types of research questions that can be answered and fail to utilize the detailed, time-specific information available from accelerometers. This study proposes the use of multilevel modeling, which tested intervention effects at specific time periods.

Design—The motivating example was the Active by Choice Today trial. Simulations were used to test whether the application of time-specific hypotheses about when physical activity intervention treatment effects were expected to occur (e.g., after-school hours) increased power to detect effects compared to traditional methods.

Methods—Six simulation conditions were tested: (1) no treatment effects (to test the type 1 error rate), (2) time-specific effects, but no traditionally-tested effects, (3) traditionally-tested effects, but no time-specific effects, and (4) combinations of traditional and time-specific effects in 3 proportions.

Results—Results showed the proposed multilevel approach demonstrated appropriate type 1 error rates and increased power to detect treatment effects during hypothesized times by 31–38 percentage points compared to traditional approaches. This was consistent across varying proportions of traditional versus time-specific effects, and there was no loss of power using the multilevel approach when only traditional effects were present.

Conclusions—The current study showed potential advantages of testing time-specific hypotheses about intervention effects using a multilevel time-specific approach. This approach may show intervention effects when traditional approaches do not. Future research should explore the application of this additional analytic tool for accelerometer physical activity estimates.

Keywords

Random coefficient models; Hierarchical linear models; CSA/MTI; Actical; Actigraph

*Corresponding author. hlawman86@gmail.com (H.G. Lawman).

1. Introduction

Accelerometer estimates of physical activity (PA) provide a potentially more accurate and flexible alternative to subjective and criterion measures.¹⁻³ Commonly used accelerometer analysis techniques include reducing the information into a single estimate of PA (e.g., average daily minutes/counts, total minutes/counts).^{4,5} This approach limits the types of research questions that can be answered and may result in reduced statistical power to find PA intervention treatment effects. While researchers have recommended a variety of analytic techniques capable of handling complex accelerometer data,^{6,7} usage and application of these techniques has been lacking.

Commonly used accelerometer analysis techniques assume PA interventions will have an effect on overall PA averaged throughout the day and across the measurement period. Descriptive studies have demonstrated differences in activity patterns across times and days of the week.⁸⁻¹⁰ It is reasonable to expect that some interventions will not have a consistent, global effect across the wear period and that researchers can make hypotheses about specific times of interest that relate to intervention goals (referred to as time-specific effects). One example is the Active by Choice Today (ACT) trial.^{11,12} ACT was a school randomized controlled trial that tested the effects of a motivation plus behavioral skills intervention (versus health education only) on increasing PA in middle school students. The ACT intervention was delivered as an after-school program and included 60 min of group PA. This suggests a specific test of the hypothesis that students in treatment schools are more active during the after-school program time periods while the intervention is in progress as well as a secondary test for whether those effects generalized to positively impact PA outside of the program time. An alternative analytic method is needed to test these hypotheses.

In circumstances where time-specific effects are applicable researchers may benefit from testing specific hypotheses about when interventions are effective. Common analytic approaches are not able to test hypotheses about treatment effects on PA patterns across time periods.⁸⁻¹⁰ Time-specific hypotheses may help researchers link intervention components to time-specific PA changes and contribute to the understanding of context and long-term behavior change.¹³ Therefore, the first goal of the current study was to demonstrate a more flexible way to utilize accelerometer data that applied time-specific hypotheses about when PA changes would occur as motivated by the published ACT outcomes.¹¹

One potential benefit of a time-specific approach is an expected increase in power to detect time-specific hypotheses about intervention effects. As traditional techniques use only a single summary score (e.g., average minutes), some changes in PA patterns across the measurement period are lost. The proposed test of time-specific hypotheses utilizes repeated accelerometer estimates of PA within each measurement period (scores during hypothesized and non-hypothesized times), using a random coefficient (multilevel model) approach.¹⁴ In effect, the entire accelerometer wear period is broken into meaningful time intervals, and treatment effects are analyzed both within and between individuals. It is expected that the within-subjects comparison should increase power over traditional between-subjects approaches.

Thus, the aims of the current study were twofold: (1) to provide a motivating example testing time-specific treatment effects, and (2) to test differences in power to assess treatment effects of the time-specific versus traditional approaches using simulations. It was hypothesized that both the time-specific and traditional methods would show appropriate type-1 error rates. Because the time-specific approach evaluates intervention effects at specific time periods it was also hypothesized that this approach would show increased power to detect treatment effects compared to the traditional approach when the intervention effects occurred only during the specified hypothesized times. The time-specific approach was not hypothesized to increase power when the intervention's true effects were the same across all time periods. In reality intervention effects may be partially time-specific and partially generalized. To better understand when this approach may be beneficial 3 additional conditions were tested in which the proportion of the effect that was time-specific varied from 33% to 66%.

2. Methods

The use of a multilevel time-specific approach was illustrated with data from the ACT trial (aim 1) in order to frame the motivating rationale for the statistical simulations (aim 2). A detailed description of the ACT trial ^{12,15} and outcomes ¹¹ have been published previously. Briefly, ACT was a one-year PA intervention involving 24 schools (12 control, 12 intervention, $n = 1431$) and was approved by the University of South Carolina Institutional Review Board. Students wore an Actical omni-directional accelerometer (Mini-Mitter, Bend, OR) for 7 days at baseline, midpoint (halfway through the academic year), and 2-weeks post-treatment. Each day of Actical data was divided into five intervals, which correspond to meaningful time blocks in the day (early morning, school hours, after school, evening, night). These intervals were 6–9am, 9am–2pm, 2–5pm, 5–8pm, and 8pm to midnight as was done in a previous national trial.¹⁶ Therefore, there were 35 PA intervals (5 time intervals per day over 7 days) per measurement period. In other applications of the proposed multilevel time-specific approach, these intervals should be customized (in both the number and length) such that they are most applicable to the research questions and intervention being tested. Raw activity data were converted into time spent in MVPA defined as ≥ 3 METS (cut-point = 1500).¹⁷

In the ACT motivating example (aim 1), both the traditional model and the time-specific model were used to examine intervention effects with sex, race, BMI, and baseline MVPA as covariates and an additional random effect to account for clustering of students within schools (see Supplement 1 for equations). MVPA during the 35 intervals at the midpoint assessment period were specified as the outcomes in the time-specific model, and it was expected that the intervention effects would be shown during the times the ACT after-school program took place (2–5pm on Monday, Tuesday, and Thursday). The ACT data were not transformed in order to retain the meaningfulness of the metric of PA minutes and are reproduced using minutes to be consistent with the published outcome data.¹¹ Secondary analyses using transformed data showed no differences.

The simulated PA data (aim 2) were modeled after ACT (35 intervals nested within child). The simulations evaluated the performance of the two analytic approaches following the

motivating example: (1) the commonly used general linear model (equivalent to an ANOVA) shown in Eq. (1) (assessing differences between treatment conditions in average MVPA) and (2) a time-specific multilevel model shown in Eq. (2) (with a dichotomous hypothesized time (HypTime) indicator coded '1' if the specific time interval is expected to show within child treatment effects and '0' otherwise, and a between child treatment effect). Equations are given below using mixed multilevel notation and included no covariates¹⁴:

$$Y_i = B_0 + B_1 TX + r_i \sim N(0, \sigma^2) \quad (1)$$

$$Y_{ti} = \gamma_{00} + \gamma_{10} \text{HypTime}_{ti} + \gamma_{01} TX_i + \gamma_{11} \text{HypTime}_{ti} TX_i + \mu_{0i} + \mu_{1i} \text{HypTime}_{ti} + r_{ti} \quad (2)$$

where: $\mu_{0i}, \mu_{1i} \sim N(0, \tau)$, $r_{ti} \sim N(0, \sigma^2)$

where TX is a treatment indicator (TX) coded 0/1. There are $t = 1, \dots, n_i$ time intervals (level-1) nested within $i = 1, \dots, I$ individuals (level-2). Using the traditional method, the effect of interest is B_1 in Eq. (1) (the average difference in MVPA between treatment and control). Using the time-specific approach in Eq. (2) there are two effects of interest: γ_{01} is the average difference in MVPA between treatment and control conditions averaged across all time points (a between person effect) and γ_{11} is the additional treatment effect specific to the hypothesized time periods versus all other times (a within person effect). In this case, the HYPtime variable was coded '1' on 2–5pm every day to reflect the applicable time-specific hypotheses and was generalized across all days (rather than the 3 ACT program days) in order to simplify the simulations. It should be noted that the proposed time-specific approach would be able to accommodate different effects across times and days consistent with the specific research questions and intervention. The traditional approach is subsumed by the time-specific approach when all time intervals are coded to show treatment effects (as hypothesized times).

For simplicity data were generated for an individually randomized rather than group randomized trial with 200 students per treatment condition to reflect sample sizes common in individually randomized interventions. For the purposes of the simulations, data were generated from a normal distribution and reflected minutes per day (or per interval) in order to facilitate the reader's understanding of effect sizes in minutes. Researchers looking to apply this analytic technique should be cautious in choosing an appropriate way for handling PA count data (e.g., transformations), which are commonly skewed.

The PA intervals were generated with a lag 1 autocorrelation of 0.3, mimicking that seen in the ACT data which results from observed PA in one interval being correlated with PA in the previous interval. PA interval means, variances, and error terms were modeled off the ACT data (see Supplement 1 for details). Error terms included: (1) the deviation of each child's intercept from the average of intercepts, (2) the deviation of each child's treatment effect at the hypothesized times from the average treatment effect at hypothesized times, and (3) the level 1 residual random error. The resulting data structure is represented in Fig. 1 as compared to a traditional summary score technique.

Six simulation conditions were run (each with 1000 replications) differing in average-weekly and time-specific treatment effects. Power was assessed by the proportion of

simulations in which the effects of interest were significant. The first simulation specified no treatment effects in order to test that the generated data exhibited appropriate type-1 error rates (i.e., a significant treatment effect in 5% of simulations). The second simulation included only a time-specific effect such that PA increased in youth in the treatment condition during the hypothesized times only, and no average-weekly treatment effect was included. The third simulation included only an average-weekly effect of treatment (no time-specific treatment effects) that mimicked current techniques which test the treatment effect on summary score PA data. The fourth, fifth, and last simulations tested varying combinations of average-weekly and time-specific treatment effects to test different effects sizes at which researchers may see benefits from using the time-specific method. Three proportions were examined: 2 and 6 (a treatment effect of 2 min per day averaged across the entire day and an additional 6 min in the hypothesized time period); 4 and 4; and 6 and 2. The data were generated to reflect a total effect of 8 min of PA mirroring that which was seen in ACT and is considered to be a small but meaningful increase.¹⁸ The R statistical package (V3.0.1) was used for all analyses and simulations,¹⁹ and full code including an example time-specific model is available in Supplement 2.

3. Results

The motivating ACT example (aim 1) showed the differences between the traditional and multilevel approaches. While both detected a significant effect of treatment at midpoint the time-specific model identified that the only significant effect occurred when students were in the after-school program (Table 1). In the traditional method, youth participating in the intervention engaged in 4.87 more minutes of MVPA during the average day compared to youth in control schools. In the time-specific method, youth participating in the intervention engaged in an average of 9.11 more minutes of MVPA during the 3 after-school program times (versus all other times) compared to youth in the control condition, but did not engage in more MVPA globally. Results showed students were engaged in MVPA during the ACT intervention at mid-point and that the intervention effects were isolated to program times, helping to explain why effects did not persist at post when the intervention had ended.

Simulation results (aim 2) are shown in Table 2. Hypothesis 1 (simulation 1) was supported in showing both methods resulted in appropriate type 1 error rates (4% and 6%) when no treatment effects were present (a null model). Hypothesis 2 (simulation 2) was supported in showing that power was higher using the time-specific method (100%) compared to the traditional method (62%) when treatment effects occurred only during specific hypothesized times (no average-weekly effects). Power was comparable across the traditional (59%) and time-specific method (57%) when there was only an average-weekly effect of treatment and no localized effects during hypothesized times (simulation 3). There was no inflation of type 1 error rates for the time-specific parameter when there was no time-specific effect in the data (5%; simulation 3). Simulations 4–6 showed that increased power to detect treatment effects persisted across the 3 combinations of average-weekly versus time-specific effects. That is, no matter the proportion of the total effect that was time-specific versus average-weekly, the time-specific method demonstrated increased power to detect treatment effects compared to the traditional method. Increases in power ranged from 31 to 38 percentages points greater for the time-specific method.

Importantly, the true type I error rate in the time-specific model should take into account both treatment parameters (treatment and treatment by time interaction). Accordingly, some adjustment for multiple comparisons is appropriate. Secondary analyses showed that the type I error rate was 0.09 when examining whether there was a significant effect for either of the treatment parameters in simulation 1 with no correction. Applying the Bonferroni correction to the treatment interaction across the simulations showed that power remained superior in the time-specific compared to the traditional approach when any of the treatment effect was localized to the hypothesized times (Table 2).

Researchers interested in translating the parameters from the time-specific method into average-weekly equivalents to facilitate comparison across studies may view an example in Supplement 1.

4. Discussion

The purpose of the current study was to demonstrate the application and to evaluate power and type-I error rates of a multilevel approach for testing time-specific treatment effects using accelerometer estimates of PA. Two key findings supported the use of this new approach.

First, the flexibility of the time-specific approach allows researchers to answer different types of research questions and test context-specific hypotheses surrounding increased PA. In the ACT trial, it was shown that students in the intervention schools (versus control) increased PA by 9.1 min during program hours or by 4.9 min when the effect was averaged throughout the day. Future research could use the time-specific approach to study generalization of PA intervention effects beyond targeted intervention times²⁰ in a wide range of settings. Some examples, among others, may include school physical education lessons²¹ which could be expected to have time-specific effects during specific PE class periods; workplace policies²² which could be expected to change PA during rather than outside work hours; or text message interventions which could be expected to increase PA in the hours after message receipt. Other applications could include momentary ecological assessments, to inform common limitations of using accelerometers in PA studies (e.g., distinguishing leisure-time and lifestyle activity, planned exercise), and a more contextual understanding of facilitators and impediments to increasing PA (e.g., time-specific targets). The proposed approach is capable of handling multiple intervention assessments (e.g., baseline, midpoint, post). In this study this was accomplished by using baseline data as a covariate. This could also be accomplished through an additional level (i.e., the use of a within subject effect across assessment points). In doing so, researchers would need to plan appropriately to ensure sufficient power, suitable sample sizes across levels, and meaningful conceptualization of levels of analysis.⁷

Researchers should be strongly cautioned that the time-specific approach relies on correctly specifying the time-specific hypotheses (choosing the correct times the intervention will have an effect). The proposed model is only applicable when there are relevant hypotheses surrounding time-specific effects, and misspecification of the model may show no added benefit or a cost when employing the time-specific approach. Furthermore, time-specific

hypotheses should be made a priori with careful attention to the theoretical and implementation aspects of the intervention. The current application can be thought of as one analytic strategy that utilizes accelerometer data in more complex ways than traditional techniques. However, additional analytic strategies are also relevant to complex uses of accelerometer data^{6,7,23} and could accommodate time-specific hypotheses. For example, growth curve modeling using a linear time variable rather than a dichotomous time effect variable may be more appropriate in cases where an intervention is expected to have a systematic linear effect across multiple assessment periods. The benefit of the time-specific approach stems from the use of a time-specific effect and a within by between interaction (rather than the models per se) and could be adapted for other models.

Secondly, results of the simulations demonstrated a strong increase (62% versus 100%) in power when the treatment effect occurred during the specific hypothesized times. The advantage in power was consistent across all ratios of generalized versus specific effects with no observed inflation of Type 1 error rates. Researchers adopting this approach should consider its use of two tests for treatment effects. Similar to other tests with multiple comparisons,²⁴ results from the current study show that some correction is appropriate. Several options are available including, treating one effect as secondary, conducting a joint test of both effects as the primary intervention test, applying unequal weighting to each effect, or other post hoc options. In developing interventions, researchers should be thoughtful about when and under what conditions their intervention would be likely to show increases in PA. As such, this more targeted approach can increase power to find an effect when traditional models would otherwise not and encourages the use of multilevel models to model between and within effects as has been suggested with other types of data.²³

Limitations of the current study should also be considered. The current study assumes the data are complete as it was outside the scope of the study to examine differential effects of missing data patterns and analytic strategies. Recommended methods for handling missing accelerometer data¹⁶ such as multiple imputation and maximum likelihood are suitable for use with the proposed model. In addition, the standardization of procedures for data reduction vary widely across studies,^{3,5} and the benefit of either novel or traditional analysis strategies is limited without greater standardization.^{5,16,25} The current study also has important strengths including simulations with high real-world applicability and the presentation of a flexible model capable of being applied to a number of accelerometer studies.

5. Conclusion

The current study supports the applicability of time-specific hypotheses and offers an additional data analytic tool for accelerometer estimated PA that is more powerful and able to take advantage of the rich accelerometer data. As seen in the motivating ACT example, research can benefit from careful consideration as to the times of day their interventions will show effects. The new multilevel time-specific approach allows researchers to find treatment effects that would otherwise be missed using traditional approaches. Future research should continue to explore and test the application of the proposed approach.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

RRP serves as a consultant and on the board for Curves, Inc. No other authors reported financial disclosures. The current study was supported by several National Institutes of Health grants (R01 HD 045693 and R01HD072153 to DKW and F32 DK100248 to HGL).

References

1. Reiser LM, Schlenk EA. Clinical use of physical activity measures. *J Am Acad Nurse Pract.* 2009; 21(2):87–94. [PubMed: 19228246]
2. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. *Sports Med.* 2001; 31(6):439–454. [PubMed: 11394563]
3. Rowlands AV. Accelerometer assessment of physical activity in children: an update. *Pediatr Exerc Sci.* 2007; 19(3):252–266. [PubMed: 18019585]
4. Esliger DW, Tremblay MS. Physical activity and inactivity profiling: the next generation. *Appl Physiol Nutr Metab.* 2007; 32:S195–S207.
5. Esliger DW, Copeland JL, Barnes JD, et al. Standardizing and optimizing the use of accelerometer data for free-living physical activity monitoring. *J Phys Act Health.* 2005; 3:366–383.
6. Hser Y, Shen H, Chou C-P, et al. Analytic approaches for assessing long-term treatment effects: examples of empirical applications and findings. *Eval Rev.* 2001; 25:233–262. [PubMed: 11317718]
7. Masse LC, Dassa C, Gauvin L, et al. Emerging measurement and statistical methods in physical activity research. *Am J Prev Med.* 2002; 23(2S):44–45. [PubMed: 12133737]
8. Jago R, Anderson CB, Baranowski T, et al. Adolescent patterns of physical activity: differences by gender, day, and time of day. *Am J Prev Med.* 2005; 28(5):447–452. [PubMed: 15894148]
9. Matthews CE, Ainsworth BE, Thompson RW, et al. Sources of variance in daily physical activity levels as measured by an accelerometer. *Med Sci Sports Exerc.* 2002; 34(8):1376–1381. [PubMed: 12165695]
10. Hesketh KR, McMinn AM, Ekelund U, et al. Objectively measured physical activity in four-year-old british children: a cross-sectional analysis of activity patterns segmented across the day. *Int J Behav Nutr Phys Act.* 2014; 11(1) Epub ahead of print.
11. Wilson DK, Van Horn ML, Kitzman-Ulrich H, et al. Results of the “active by choice today” (act) randomized trial for increasing physical activity in low income and minority adolescents. *Health Psychol.* 2011; 30(4):463–471. [PubMed: 21534677]
12. Wilson DK, Kitzman-Ulrich H, Williams JE, et al. An overview of “the active by choice today” (act) trial for increasing physical activity. *Contemp Clin Trials.* 2008; 29(1):21–31. [PubMed: 17716952]
13. Deci EL, Ryan RM. Facilitating optimal motivation and psychological well-being across life’s domains. *Can Psychol.* 2008; 49(1):14–23.
14. Raudenbush, SW.; Bryk, AS. Hierarchical linear models: applications and data analysis methods. Thousand Oaks, CA: Sage Publications Inc; 2002.
15. Wilson DK, Griffin S, Saunders R, et al. Using process evaluation for program improvement in dose, fidelity, and reach: the act experience. *Int J Behav Nutr Phys Act.* 2009; 6(79)
16. Catellier DJ, Hannan PJ, Murray DM, et al. Imputation of missing data when measuring physical activity by accelerometry. *Med Sci Sports Exerc.* 2005; 37:S555–S562. [PubMed: 16294118]
17. Puyau MR, Adolph AL, Vohra FA, et al. Prediction of activity energy expenditure using accelerometers in children. *Med Sci Sports Exerc.* 2004; 36(9):1625–1631. [PubMed: 15354047]
18. Troiano RP, Berrigan D, Dodd KW, et al. Physical activity in the united states measured by accelerometer. *Med Sci Sports Exerc.* 2008; 40(1):181–188. [PubMed: 18091006]

19. R Core Development Team. A language and environment for statistical computing. 2013
20. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol*. 2000; 55(1):68–78. [PubMed: 11392867]
21. McKenzie TL, Sallis JF, Prochaska JJ, et al. Evaluation of a two-year middle-school physical education intervention: M-span. *Med Sci Sports Exerc*. 2004; 36(8):1382–1388. [PubMed: 15292747]
22. Crespo NC, Sallis JF, Conway TL, et al. Worksite physical activity policies and environments in relation to employee physical activity. *Am J Health Promot*. 2011; 25(4):264–271. [PubMed: 21361812]
23. Schwartz J, Stone AA. Strategies for analyzing ecological momentary assessment data. *Health Psychol*. 1998; 17(1):6–16. [PubMed: 9459065]
24. Shaffer JP. Multiple hypothesis testing. *Annu Rev Psychol*. 1995; 46:561–584.
25. Stone MR, Rowlands AV, Eston RG. Relationships between accelerometer-assessed physical activity and health in children: impact of the activity-intensity classification method. *J Sports Sci Med*. 2009; 8(1):136–143. [PubMed: 24150567]

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi: 10.1016/j.jsams.2014.09.003.

6. Practical implications

- Researchers can use a multilevel time-specific approach to model accelerometer-estimated physical activity treatment effects that would otherwise be missed using traditional analytic approaches.
- Researchers can explore new types of research questions that target time-specific differences in accelerometer-estimated physical activity patterns and interventions.
- The new analytic approach may help clinicians develop interventions that capitalize on time-specific facilitators and strategically avoid time-specific barriers to increasing physical activity.
- The new analytic approach allows researchers to increase power to find intervention effects through careful consideration as to the times of day their interventions will show effects.

Traditional Analytic Method

ID	TX	MVPA.Sum	Mon.Time1	Mon.Time2	Mon.Time3	Mon.Time4	Mon.Time5
1000	0	20.34	3.61	3.72	1.00	9.00	3.00
1001	1	40.35	11.67	6.80	14.23	0.76	6.89
1002	1	92.08	61.67	10.20	14.00	4.53	1.68

$$MVPA.Sum_i = B_0 + B_1TX + r_i$$

Mon.Time1 – Mon.Time5 are traditionally summed to create MVPA.Sum and discarded for analyses.

New Analytic Method

ID	TX	MVPA.Sum	Interval	MVPA.Interval	HypTime
1000	0	20.34	1	3.61	0
1000	0	20.34	2	3.72	0
1000	0	20.34	3	1.00	1
1000	0	20.34	4	9.00	1
1000	0	20.34	5	3.00	0
1001	1	40.35	1	11.67	0
1001	1	40.35	2	6.80	0
1001	1	40.35	3	14.23	1
1001	1	40.35	4	0.76	1
1001	1	40.35	5	6.89	0
1002	1	92.08	1	61.67	0
1002	1	92.08	2	10.20	0
1002	1	92.08	3	14.00	1
1002	1	92.08	4	4.53	1
1002	1	92.08	5	1.68	0

$$MVPA.Interval_{ti} = \gamma_{00} + \gamma_{10}HypTime_{ti} + \gamma_{01}TX_i + \gamma_{11}HypTime_{ti}TX_i + \mu_{0i} + \mu_{1i}HypTime_{ti} + r_{ti}$$

The new method uses all the MVPA block variables in a longitudinal approach and adds a variable that indicates which MVPA Blocks the treatment is hypothesized to have its effect.

Fig. 1.

An illustration of the data structure using the traditional analytic technique versus the new analytic technique.

Table 1

An applied example of time-specific hypotheses for treatment effects at midpoint versus the traditional summary score methodology in the Active by Choice Today (ACT) trial.

Parameter	Time-specific analytic technique				Traditional analytic technique			
	Estimate	Lower CI	Upper CI	df	mi	Estimate	Lower CI	Upper CI
Intercept	9.05*	7.50	10.60	319	0.29	44.49*	51.37	37.61
Female	-2.06*	-2.67	-1.45	372	0.26	-9.78*	-7.15	-12.41
Black	-0.55	-1.24	0.14	350	0.27	-1.70	1.18	-4.58
Baseline BMI	-0.13*	-0.17	-0.09	644	0.17	-0.63*	-0.41	-0.85
Baseline MVPA	0.06*	0.04	0.08	98	0.58	0.29*	0.35	0.23
TX	0.25	-0.44	0.94	22	0.16	4.87*	8.36	1.38
Program Time	2.30*	4.59	0.01	109	0.08	-	-	-
Program Time*TX	9.11*	12.30	5.92	22	0.06	-	-	-

Note: (se) is the standard error of the parameter estimate; adjusted for the use of multiple imputation; df is the estimated df adjusted for use of multiple imputation; mi is the fraction of missing information, estimates with $p < 0.05$ are indicated with an *. Analyses reprinted in part from Wilson et al.¹¹

Simulation results testing the application of time-specific hypotheses about treatment effects compared to traditional summary score techniques.

Table 2

	Traditional analytic method				New analytic method			
	B	Lower CI	Upper CI	Power	B	Lower CI	Upper CI	Power
Simulation 1: no treatment effects, null model								
Intercept	43.52	38.55	48.19	1.00	8.70	7.70	9.63	1.00
Treatment	0.02	-6.30	7.12	0.04	0.01	-1.30	1.40	0.04
Hypothesized times	-	-	-	-	0.02	-0.82	0.82	0.06
Treatment by Hyp. Time	-	-	-	-	-0.01	-1.20	1.19	0.06
Simulation 2: no average-weekly effect, 8 min time-specific effect								
Intercept	43.55	38.55	48.42	1.00	8.71	7.69	9.67	1.00
Treatment	7.99	1.31	14.69	0.62	-0.01	-1.38	1.33	0.05
Hypothesized times	-	-	-	-	-0.01	-0.82	0.82	0.05
Treatment by Hyp. Time	-	-	-	-	8.01	6.90	9.24	1.00
Simulation 3: 8 min average-weekly effect only, no time-specific effect								
Intercept	43.69	38.70	48.57	1.00	8.74	7.73	9.70	1.00
Treatment	7.72	0.64	14.63	0.59	1.55	0.19	2.96	0.57
Hypothesized times	-	-	-	-	0.01	-0.80	0.82	0.05
Treatment by Hyp. Time	-	-	-	-	-0.02	-1.11	1.13	0.05
Simulation 4: 2 min average-weekly effect, 6 min time-specific effect								
Intercept	43.43	38.64	48.45	1.00	8.69	7.73	9.72	1.00
Treatment	8.06	1.69	14.81	0.62	0.41	-0.92	1.76	0.08
Hypothesized times	-	-	-	-	-0.02	-0.86	0.79	0.05
Treatment by Hyp. Time	-	-	-	-	6.02	4.90	7.17	1.00
Simulation 5: 4 min average-weekly effect, 4 min time-specific effect								
Intercept	43.44	38.44	48.29	1.00	8.69	7.70	9.65	1.00
Treatment	8.26	0.92	14.62	0.66	0.86	-0.56	2.15	0.21
Hypothesized times	-	-	-	-	0.01	-0.79	0.82	0.05
Treatment by Hyp. Time	-	-	-	-	3.98	2.87	5.09	1.00
Simulation 6: 6 min average-weekly effect, 2 min time-specific effect								
Intercept	43.49	38.42	48.26	1.00	8.70	7.67	9.68	1.00

	Traditional analytic method				New analytic method			
	<i>B</i>	Lower CI	Upper CI	Power	<i>B</i>	Lower CI	Upper CI	Power
Treatment	8.01	1.41	14.95	0.61	1.20	-0.12	2.59	0.38
Hypothesized times	-	-	-	-	0.00	-0.84	0.82	0.05
Treatment by Hyp. Time	-	-	-	-	1.99	0.79	3.15	0.92

CI = 95% confidence intervals. *Note:* estimates were averaged across 1000 simulations. Confidence intervals were obtained by taking the 2.5% and 97.5% of the distribution of estimates. Power was calculated as the proportion of significant simulations. Bonferroni corrected power estimates for the treatment interaction parameter were 0.04, 0.02, 1.00, 1.00, 1.00, and 0.88, respectively, across simulations.