

Individual Information-Centered Approach for Handling Physical Activity Missing Data

Minsoo Kang, David A. Rowe, Tiago V. Barreira, Terrance S. Robinson, and Matthew T. Mahar

The purpose of this study was to validate individual information (II)-centered methods for handling missing data, using data samples of 118 middle-aged adults and 91 older adults equipped with Yamax SW-200 pedometers and Actigraph accelerometers for 7 days. We used a semisimulation approach to create six data sets: three physical activity outcome measurements (i.e., step counts, activity counts, and minutes of moderate to vigorous physical activity) for both groups (i.e., middle-aged adults and older adults). After analyzing each data set separately, we replaced missing values with two II-centered and two group information (GI)-centered methods. Root mean square difference (RMSD), mean signed difference, paired t tests, and Pearson correlations were used to determine the effectiveness of the various recovery methods. Overall, the II-centered methods showed smaller RMSDs than the GI-centered methods for each data set in both groups. We found no significant mean differences between the known values and the replacement values in all conditions. The II-centered methods produced better results than GI-centered methods. We determined substituting missing data points using the average of days remaining to be an accurate missing data recovery method for middle-aged adults' and older adults' pedometer and accelerometer data.

Key words: accelerometer, missing value, pedometer, recovery method

In studying the relationship of physical activity to health, researchers have introduced several objective field measures to assess physical activity behavior (Haskell et al., 2007; Nelson et al., 2007; Dale, Welk, & Matthews, 2002). Pedometers and accelerometers are popular tools because they are small, unobtrusive, and easy to use (Bassett & Strath, 2002; Welk, 2002). Pedometers track the number of steps a person walks and can be used to measure an individual's physical activity level and changes in physical activity patterns during intervention studies (Chan, Ryan, & Tudor-Locke, 2004; Tudor-Locke et al., 2004; Tudor-Locke, Williams, Reis, & Pluto, 2002). Accelerometers measure activity counts and minutes of

moderate to vigorous physical activity (MVPA) and are used to assess the duration and intensity of physical activity during observational (Hagstromer, Oja, & Sjostrom, 2007) and experimental studies (Meijer, Westerterp, & Verstappen, 1999; Sloomaker, Chin A Paw, Schuit, Seidell, & van Mechelen, 2005).

Accurately measuring physical activity behavior, however, is challenging due to high inter- and intraindividual variability. Variables such as weather, season, or day of the week can produce disparate results (Chan, Ryan, & Tudor-Locke, 2006; Matthews, Ainsworth, Thompson, & Bassett, 2002; Tudor-Locke et al., 2004). Because of the high variability, researchers have tried to determine more appropriate data collection periods. Gretebeck and Montoye (1992) reported that at least 5 or 6 consecutive days of adult pedometer data collection are needed to minimize individual variability to an acceptable degree. They also recommended that both weekdays and weekend days be included. Tudor-Locke et al. (2005) concluded that a minimum of 3 days of adult pedometer data is sufficient to estimate free-living physical activity in a week. Matthews et al. (2002) suggested that 7 days of accelerometer data collection are needed for a stable estimation of adults' physical activity patterns. Because stable measures of physical activity often require data collection on multiple days, missing values are a common phenomenon.

Submitted: October 16, 2007

Accepted: March 2, 2008

Minsoo Kang and Tiago V. Barreira are with the Department of Health and Human Performance at Middle Tennessee State University. David A. Rowe is with the Department of Sport, Culture, and the Arts at the University of Strathclyde. Terrance S. Robinson and Matthew T. Mahar are with the Department of Exercise and Sport Science at East Carolina University.

Missing data can threaten the validity of a research study (Roth, 1994). Important information can be lost if there is a large amount of incomplete or missing data. If cases with missing values are not included in the study, the data analysis can be compromised due to a decrease in total sample size, which results in less statistical power. Additionally, missing data have a negative effect on the quality of research, forcing researchers to spend more time and money recruiting larger sample sizes. To eliminate or reduce the threat of missing data, statistical methods have been developed to recover the missing values.

The most common recovery method is GI, in which a summary (e.g., mean) from the sample group replaces an individual's missing value (Acock, 1997; Little & Rubin, 1989). The GI-centered approach, however, may not be appropriate in handling step count data when repeated measures are used (i.e., multiple days of data collection). Replacements based on intraindividual information may generate a more accurate recovery of the missing values because intraindividual variability is lower than interindividual variability (Tudor-Locke et al., 2005). Laird (1988) pointed out that when data are collected repeatedly over time on the same experimental units (e.g., an individual) the use of the GI-centered method is inappropriate, results in a loss of efficiency, and may bias the results. Schafer and Graham (2002) recommended procedures that use all available data for each participant. They believed that missing information could be partially recovered from the earlier or later data from the same individual. Therefore, using the II-centered method should enable an effective recovery of the missing values.

Researchers have attempted to handle physical activity missing data using II. In 2003, Kang and Zhu (2003) introduced an II-centered approach in recovering missing values for step count data. They compared several missing data recovery methods and concluded that II-centered methods are superior to GI-centered techniques, including regression imputations and expectation maximization. Rowe, Mahar, Raedeke, and Lore (2004) used the II-centered method in studying children's physical activity behaviors. In addition, Kang, Zhu, Tudor-Locke, and Ainsworth (2005) compared several conditions of weekdays and/or weekend days in pedometer step counts using II- and GI-centered methods and found that II-centered methods produced better data recovery. The II-centered approach for handling missing values in step counts has potential, but researchers should consider certain factors before making generalizations.

In the Kang et al. (2005) study, participants wore pedometers during waking hours and recorded their step counts for 21 consecutive days with a random starting date. From 21 days of data collection, the authors were able to use eight different conditions for data replacement, but the normal practice for collecting pedometer and accelerometer data is between 3 and 7 days, limiting the

number of conditions that can be used. Use of the II-centered method in handling missing data for 7-day studies should be validated. In addition, participants in the 2005 study ranged in age from 17 to 79 years old. Using specific populations (e.g., middle-aged adults and older adults) may contribute to different results. Another important question is how II-centered methods would function when applied to other physical activity measurement tools, such as accelerometers. Replacement methods by imputation can be successfully applied to accelerometer data (Catellier et al., 2005), but due to the complexity of those methods (i.e., expectation maximization and multiple imputation) and the requirements of large sample sizes, an investigation of II-centered methods to replace missing accelerometer data is warranted. Thus, the purpose of our study was to validate II-centered methods for handling missing data with specific samples of middle-aged adults' and older adults' pedometer and accelerometer data sets during 7 days of data collection.

Method

All procedures received institutional review board approval, and participants' informed consent was obtained before data were collected.

Participants

The middle-aged participant group was composed of 118 adults between 40 and 60 years of age. The older participant group was composed of 91 adults over 60 years of age. Participants were recruited from local communities via newspaper advertisements and direct contact with local agencies, such as church groups, community centers, assisted living centers, and recreational fitness groups. The middle-aged participants were studied in eastern North Carolina (65% local), West Virginia (29%), and four other states (6%). The older adults were recruited mostly from eastern North Carolina communities (80%). A subsample (20%) was recruited from West Virginia. Participants were instructed to wear the instruments for 7 complete days. When data collection was complete, they returned the instruments in postage-paid envelopes.

Physical Activity Outcome Measurements

Step counts were measured with pedometers (SW-200; Yamax Corp., Tokyo, Japan), and activity counts were measured with accelerometers (Model 7164; Actigraph LLC, Fort Walton Beach, FL). Minutes of MVPA were calculated using the activity counts above a cut point of 1,951 counts/minute, established by Freedson, Melanson, and Sirard (1998), for moderate intensity physical activity.

Semisimulation Data Generation

Traditional ways to simulate data to study the effectiveness of missing data recovery methods (e.g., Myers, 2000) may not be useful in practical data replacement situations because the data characteristics often do not take into consideration the true pattern of missing values. Our study used a semisimulation design to create missing data sets. The semisimulation approach modifies an existing data set by incorporating the actual pattern of missing data to determine the most effective recovery method (Kang et al., 2005).

First, we selected participants with complete data to form a nonmissing data sample. After screening the remaining participants, we kept only those with at least 1 weekend day and 3 weekdays to form a missing data sample. This exclusion criterion was developed, as large amounts of missing data may reduce the accuracy of recovery (Roth, 1994; Stumpf, 1978). The missing data pattern was copied, and a semisimulation data sample was created in a one-to-one fashion using the nonmissing data.

We created semisimulation data sets using similar procedures for each outcome variable. Procedures for the step count data are provided to illustrate the procedure. Ninety-six of the 118 middle-aged adults had no missing data, and the remaining 22 had at least one missing data point. We screened the missing data from 22 participants following previously described guidelines, and kept a total of 16 participants with missing values. To maintain the same pattern as the real data, the missing data sample ($n = 16$) with truly missing values was copied to the nonmissing data sample ($n = 96$) until every participant had at least one missing data point. Then we created a semisimulated missing data sample. Because of the different physical activity outcome measurements and samples, we created a total of six data sets (i.e., 3 Outcome Measurements \times 2 Groups; see Tables 1 and 2). Kang et al. (2005) detailed the procedure for creating semisimulation data sets.

Recovery Method

Mean substitution methods were used in the study (Allison, 2001). Each individual's average was used for II-centered methods, and the group's average was used for GI-centered methods. We examined several conditions, including combinations of weekday and weekend day information depending on the type of missing day. The created missing values were replaced with the following four conditions:

Individual information (II)-centered:

- Mean of remaining days of the participant
- Mean of participant's remaining weekdays or weekends, depending on the type of missing day

Group information (GI)-centered:

- Mean of same day (e.g., Monday) of all participants
- Mean of type of day of all participants (e.g., weekday)

Data Analysis

We evaluated the effect of various combinations in recovering missing values by comparing the known values purposely removed from the data set with the replacement values estimated by the four different recovery conditions. RMSD and mean signed difference (MSD) were used to determine the effectiveness of the various missing data recovery methods (Catellier et al., 2005; Kang et al., 2005). RMSD was calculated by the differences between the original and replacement values, which were then squared, averaged, and square-rooted. The formula for RMSD is as follows, where $Nmvp$ refers to the total number of missing value points:

$$RMSD = \sqrt{\frac{\sum_{j=1}^{Nmvp} (\text{original value}_j - \text{replacement value}_j)^2}{Nmvp}} \quad (1)$$

MSD was calculated by the differences between the original and replacement values, which were then averaged. The formula for MSD follows:

$$MSD = \frac{\sum_{j=1}^{Nmvp} (\text{original value}_j - \text{replacement value}_j)}{Nmvp} \quad (2)$$

A smaller RMSD and close-to-zero MSD represent a better recovery of the missing values. Paired t tests were used to examine the mean differences between the original and the replacement values. We used the Bonferroni technique to adjust the alpha level. In addition, to determine whether the data replacement methods altered relationships with other variables, we calculated Pearson correlation coefficients between the original data set and health-related variables (i.e., body mass index [BMI], daily minutes of MVPA), and between the "generated" data set and the health-related variables. We used Microsoft Excel and SPSS 15.0 statistical software to complete the analyses.

Results

Missing Data Recovery

The RMSD and MSD results for both II- and GI-centered methods are summarized in Table 3. A smaller RMSD value represents a better recovery of the missing values. The RMSD for middle-aged adults ranged from 4,444 to 5,296 for step counts, 128,037 to 168,496 for activity counts, and 21.8 to 29.7 for minutes of MVPA. The RMSD for older adults ranged from 1,620 to 3,790 for step counts, 74,005 to 111,671 for activity counts, and 14.8 to 21.4 for minutes of MVPA. Overall, the II-centered methods showed smaller RMSDs than the GI-centered methods, and results from the older adults' data set showed smaller RMSDs than those from the middle-aged adults' data. The differences in the RMSD between the II- and GI-centered methods were consistently less for middle-aged adults than for older adults, across all three

measures. The II-centered method, more specifically condition “a” (i.e., replacing a missing value by the mean of the participant’s remaining days), was most effective in recovering the missing values for all measures.

MSD was included to show the direction and degree of bias that may be caused by missing data recovery methods. Unlike the RMSD results, we found no condition to be superior. Negative MSDs across some conditions indicated predicted missing values tended to be overestimated. Paired *t* test results are also presented under each condition in Table 3. We found no statistically significant mean differences between the original and the replacement values in all conditions. The Pearson correlation coefficients are presented in Table 4. Overall, the correlations between the original data and health measures, and the replacement condition “a” and health measures, were similar, with the differences ranging from .00 to .03.

Discussion

Missing data can negatively affect physical activity research studies, and a practical solution is needed. Missing data can be restored to minimize information loss and bias. In our study, we found the II-centered methods to be superior to the GI-centered methods for handling missing data from pedometers (i.e., step counts) and from ac-

celerometers (i.e., activity counts and minutes of MVPA) for both middle-aged adults and older adults. Kang et al. (2005) also found that the II-centered methods were superior to GI-centered methods for pedometer step count data. The results not only demonstrated that the II-centered approach for replacing missing pedometer data works with different homogeneous samples, but found the II-centered approach to be a viable alternative for replacement of missing accelerometer data. In contrast to the results from Kang et al. (2005), the replacement condition that produced the lowest RMSD was the II-centered method with the mean of remaining days, and not the II-centered method with the mean of remaining weekdays or weekend days, depending on the type of missing day. This disparity could be due to the fewer days of data collection for the current study data. When the II-centered method with the mean of remaining weekdays or weekend days was applied, only one weekend day was available to replace the missing value, which eliminated any possible variation of the measure for the weekend.

One challenge in measuring physical activity behavior is its high inter- and intraindividual variability. To account for the variability, different replacement methods were analyzed, but the number of conditions was limited compared to the previous study (Kang et al., 2005). The difference in the number of replacement conditions occurred due to number of days of data collection: 21 days (6 weekend days) in the study by Kang et al. (2005) and 7 days (2 weekend days) in our study. The number of days of data collection and the conditions applied in our study were closer to the normal practice.

The RMSDs showed a wide range between the different outcome measurements, ranging from 14.8 for older adults’ minutes of MVPA to 168,496 for middle-aged adults’ activity counts. Because RMSDs depend on the unit of measurement, the RMSDs of the various outcome variables are not directly comparable. RMSDs also seem to be dependent on the score magnitude. The RMSDs for middle-aged adults were higher than those for older adults, and the middle-aged adults’ scores showed greater variation than older adults in all measurements, as seen in Table 3. Regardless of outcome measurement or group,

Table 1. Data sets and sample sizes

Outcome measurements	Groups	Total (N)	Nonmissing (n)	Missing (n)
Step counts	Middle-aged adults	118	96	16
	Older adults	91	75	15
Activity counts and MVPA	Middle-aged adults	118	58	36
	Older adults	91	62	9

Note. MVPA = minutes of moderate to vigorous physical activity.

Table 2. Semisimulation data set: sample size and missing value

Outcome measurements	Groups	Semisimulated (n)	Missing 1 day (n)	Missing 2 or more days (n)	Total missing value points (N_{mvp})	Percentage of missing values
Step counts	Middle-aged adults	96	60	36	144	21.4
	Older adults	75	60	15	90	17.1
Activity counts and MVPA	Middle-aged adults	58	36	22	84	20.7
	Older adults	62	27	35	104	24.0

Note. MVPA = minutes of moderate to vigorous physical activity.

the II-centered method with replacement by the mean of remaining days showed the lowest RMSD.

Researchers have introduced more sophisticated approaches to handling missing values with repeated

physical activity measures (Catellier et al., 2005; Kang & Zhu, 2003). Those approaches include expectation maximization (Little & Rubin, 1987) and multiple imputation (Schafer, 1997). These iterative methods may provide

Table 3. Recovery statistics by physical activity measurements and groups

Outcome measurements	Groups	Statistical index	Conditions			
			a	b	c	d
Step counts	Middle-aged adults	RMSD	4,444	4,549	5,297	5,255
		MSD	434	627	841	689
		<i>t</i> test	1.17	1.67	1.92	1.58
	Older adults	RMSD	1,620	1,898	3,790	3,610
		MSD	-214	-272	287	69
		<i>t</i> test	-1.26	-1.37	0.72	0.18
Activity counts	Middle-aged adults	RMSD	128,037	147,637	168,496	166,203
		MSD	-10,715	2,178	322	-6,582
		<i>t</i> test	-0.77	0.13	0.02	-0.36
	Older adults	RMSD	74,005	75,779	111,632	111,671
		MSD	4,981	-4,080	6,964	6,841
		<i>t</i> test	0.69	-0.55	0.63	0.62
MVPA	Middle-aged adults	RMSD	21.8	25.3	29.6	29.7
		MSD	-2.4	1.1	-1.0	-1.0
		<i>t</i> test	-1.01	0.39	-0.32	-0.32
	Older adults	RMSD	14.8	15.0	21.3	21.4
		MSD	-0.5	-2.4	-1.2	-1.3
		<i>t</i> test	-0.36	-1.63	-0.56	-0.63

Note. The *t* tests are between the original and replacement values; the II-centered methods use the mean of: (a) = the remaining days and (b) = the remaining weekdays or weekend days, depending on the type of missing day; the GI-centered methods use the mean of: (c) = the same day that missing value occurred and (d) = the weekdays or weekend days depending on the type of missing day; MVPA = minutes of moderate to vigorous physical activity; RMSD = root mean square difference; MSD = mean signed difference.

Table 4. Pearson correlation coefficients by physical activity measurements and groups

Outcome measurements	Groups	Health measurements	Original	Conditions			
				a	b	c	d
Step counts	Middle-aged adults	BMI	-.39	-.38	-.39	-.39	-.39
		MVPA	.57	.60	.61	.62	.62
	Older adults	BMI	-.19	-.19	-.19	-.21	-.20
		MVPA	.81	.81	.81	.82	.82
Activity counts	Middle-aged adults	BMI	-.25	-.24	-.24	-.24	-.24
		Step counts	.70	.70	.71	.69	.69
	Older adults	BMI	-.22	-.19	-.20	-.18	-.18
		Step counts	.80	.80	.79	.78	.78
MVPA	Middle-aged adults	BMI	-.16	-.18	-.18	-.17	-.17
		Step counts	.57	.57	.58	.56	.56
	Older adults	BMI	-.34	-.34	-.34	-.32	-.32
		Step counts	.80	.80	.79	.78	.79

Note. Original = nonmissing sample; the II-centered methods use the mean of: (a) = the remaining days and (b) = the remaining weekdays or weekend days, depending on the type of missing day; the GI-centered methods use the mean of: (c) = the same day that missing value occurred and (d) = the weekdays or weekend days depending on the type of missing day; BMI = body mass index; MVPA = minutes of moderate to vigorous physical activity.

unbiased estimates of physical activity data collected over time, but they have a number of shortcomings including the complexity of the methods and the requirements of large sample sizes. In addition, a statistical method, mixed modeling with estimation by maximum likelihood, was suggested for longitudinal data sets (Laird & Ware, 1982). Mixed modeling adjusts for missing values by accounting for all fixed and random effects in the model. However, these approaches are often not practical for researchers who may not have the expertise or software to use such methods. Rather, the II-centered approach is relatively straightforward and more accurate, compared to the expectation maximization method (Kang & Zhu, 2003).

Using group mean substitution methods can cause a loss in the data set variance (Allison, 2001; Roth, 1994). Descriptive statistics, including mean and standard deviation, are shown in Table 5. As expected, there was a loss of variance for the GI-centered methods. The II-centered methods, however, resulted in variance similar to the original data set. Thus, the II-centered methods preserved the variance more than the traditional GI-centered methods. Results from the relationships with other health-related variables between the original data set and the data set with replaced values show another strength of the II-centered approach to replace missing data. The correlation coefficients (see Table 4) between the “generated” data based on the II-centered methods and health variables were similar to those between the original data and health variables. There were similar results for every outcome variable and both age groups.

A limitation of using RMSD is that there is no criterion for an absolute decision for physical activity measure-

ments (Kang et al., 2005), making it difficult to compare results across studies or across populations within the same study. But by using *t* tests, we determined the replaced values were not significantly different from the original values. In addition, the correlations between BMI and values from the original data set and from the data set with replacement values were similar. This indicates that II-centered data replacement methods do not affect the relationship with other variables.

Another issue in our study was the amount of missing values created in the semisimulated data set, which may not truly represent a typical missing data situation. It is difficult to determine what a typical missing data percentage is in studies using pedometers and accelerometers. The percentage of missing values created in our study varied from 17.1 to 24% (see Table 2), which were the same percentages found in the missing sample data sets. In addition, missing data may bias the remaining data. For example, less active individuals may be less likely to use the pedometers or accelerometers because they do not want to reveal their inactivity levels.

Determining an effective method to recover missing data can be valuable for physical activity researchers. An appropriate data recovery method can help to reduce the time and cost of data collection, increase the quality of the data, and prevent a loss of statistical power. We demonstrated that the II-centered methods are superior to the GI-centered methods for handling pedometer and accelerometer missing data when 7 days of data are collected. The II-centered methods produced data sets similar to the original data sets and similar correlations with health-related variables.

Table 5. Statistical summary by physical activity measurements and groups

Outcome measurements	Groups	Statistical index	Conditions				
			Original	a	b	c	d
Step counts	Middle-aged adults	<i>M</i>	9,684	9,591	9,550	9,504	9,536
		<i>SD</i>	3,788	3,866	3,848	3,095	3,091
	Older adults	<i>M</i>	4,996	5,032	5,042	4,946	4,984
		<i>SD</i>	3,315	3,363	3,350	2,688	2,716
Activity counts	Middle-aged adults	<i>M</i>	319,114	321,330	318,663	319,047	320,475
		<i>SD</i>	120,564	122,574	118,246	94,903	95,299
	Older adults	<i>M</i>	203,697	202,504	204,675	202,028	202,058
		<i>SD</i>	99,461	101,337	102,647	78,721	78,960
MVPA	Middle-aged adults	<i>M</i>	29.8	30.3	29.6	30.0	30.0
		<i>SD</i>	21.6	21.7	20.8	16.8	16.8
	Older adults	<i>M</i>	15.0	15.1	15.6	15.3	15.3
		<i>SD</i>	17.3	17.6	17.7	13.5	13.5

Note. Original = nonmissing sample; the II-centered methods use the mean of: (a) = the remaining days and (b) = the remaining weekdays or weekend days, depending on the type of missing day; the GI-centered methods use the mean of: (c) = the same day that missing value occurred and (d) = the weekdays or weekend days depending on the type of missing day; *M* = mean; *SD* = standard deviation; MVPA = minutes of moderate to vigorous physical activity.

References

- Acock, A. C. (1997). Working with missing values. *Family Science Review*, 10(1), 76–102.
- Allison, P. D. (2001). *Missing data* (Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-136, pp. 11–12). Thousand Oaks, CA: Sage.
- Bassett, D. R., Jr., & Strath, S. J. (2002). Use of pedometers to assess physical activity. In G. J. Welk (Ed.), *Physical activity assessments for health-related research* (pp. 163–177). Champaign, IL: Human Kinetics.
- Catellier, D. J., Hannan, P. J., Murray, D. M., Addy, C. L., Conway, T. L., Yang, S., et al. (2005). Imputation of missing data when measuring physical activity by accelerometry. *Medicine & Science in Sports & Exercise*, 37(Suppl 11), S555–S562.
- Chan, C. B., Ryan, D. A. J., & Tudor-Locke, C. (2004). Health benefits of a pedometer-based physical activity intervention in sedentary workers. *Preventive Medicine*, 39, 1215–1222.
- Chan, C. B., Ryan, D. A. J., & Tudor-Locke, C. (2006). Relationship between objective measures of physical activity and weather: A longitudinal study. *International Journal of Behavioral Nutrition and Physical Activity*, 3, 21.
- Dale, D., Welk, G. J., & Matthews, C. E. (2002). Methods for assessing physical activity and challenges for research. In G. J. Welk (Ed.), *Physical activity assessments for health-related research* (pp. 19–34). Champaign, IL: Human Kinetics.
- Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine & Science in Sports & Exercise*, 30, 777–781.
- Gretebeck, R., & Montoye, H. (1992). Variability of some objective measures of physical activity. *Medicine & Science in Sports & Exercise*, 24, 1167–1172.
- Hagstromer, M., Oja, P., & Sjostrom, M. (2007). Physical activity and inactivity in an adult population assessed by accelerometry. *Medicine & Science in Sports & Exercise*, 39, 1502–1508.
- Haskell, W. L., Lee, I. M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., et al. (2007). Physical activity and public health: Updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Medicine & Science in Sports & Exercise*, 39, 1423–1434.
- Kang, M., & Zhu, W. (2003). Current issues with missing data methods in physical activity research. *2003 Daegu University Conference Proceedings: Facing the Challenge* (pp. 610–616). Daegu, Korea.
- Kang, M., Zhu, W., Tudor-Locke, C., & Ainsworth, B. (2005). Experimental determination of effectiveness of an individual information-centered approach in recovering step-count missing data. *Measurement in Physical Education and Exercise Science*, 9, 233–250.
- Laird, N. M. (1988). Missing data in longitudinal studies. *Statistics in Medicine*, 7, 305–315.
- Laird, N. M., & Ware, J. H. (1982). Random effects models for longitudinal data: An overview of recent results. *Biometrics*, 38, 963–974.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Little, R. J. A., & Rubin, D. B. (1989). The analysis of social science data with missing values. *Sociological Methods and Research*, 18, 292–326.
- Matthews, C. E., Ainsworth, B. E., Thompson, R. W., & Bassett, D. R., Jr. (2002). Sources of variance in daily physical activity levels as measured by an accelerometer. *Medicine & Science in Sports & Exercise*, 34, 1376–1381.
- Meijer, E. P., Westerterp, K. R., & Verstappen, F. T. (1999). Effect of exercise training on total daily physical activity in elderly humans. *European Journal of Applied Physiology and Occupational Physiology*, 80, 16–21.
- Myers, W. R. (2000). Handling missing data in clinical trials: An overview. *Drug Information Journal*, 34, 525–533.
- Nelson, M. E., Rejeski, W. J., Blair, S. N., Duncan, P. W., Judge, J. O., King, A. C., et al. (2007). Physical activity and public health in older adults: Recommendation from the American College of Sports Medicine and the American Heart Association. *Medicine & Science in Sports & Exercise*, 39, 1435–1445.
- Roth, P. L. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology*, 47, 537–560.
- Rowe, D. A., Mahar, M. T., Raedeker, T. D., & Lore, J. (2004). Measuring physical activity in children with pedometers: Reliability, reactivity, and replacement of missing data. *Pediatric Exercise Science*, 16, 343–354.
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman and Hall.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.
- Slotmaker, S. M., Chin A Paw, M. J. M., Schuit, A. J., Seidell, J. C., & van Mechelen, W. (2005). Promoting physical activity using an activity monitor and a tailored web-based advice: Design of a randomized controlled trial. *BMC Public Health*, 5, 134.
- Stumpf, S. A. (1978). A note on handling missing data. *Journal of Management*, 4, 65–73.
- Tudor-Locke, C., Bassett, D. R., Jr., Swartz, A. M., Strath, S. J., Parr, B. B., Reis, J. P., et al. (2004). A preliminary study of one year of pedometer self-monitoring. *Annals of Behavioral Medicine*, 28, 158–162.
- Tudor-Locke, C., Burkett, L., Reis, J. P., Ainsworth, B. E., Macera, C. A., & Wilson, D. K. (2005). How many days of pedometer monitoring predict weekly physical activity in adults? *Preventive Medicine*, 40, 293–298.
- Tudor-Locke, C., Williams, J. E., Reis, J. P., & Pluto, D. (2002). Utility of pedometers for assessing physical activity. *Sports Medicine*, 32, 795–808.
- Welk, G. J. (2002). Use of accelerometry-based activity monitors to assess physical activity. In G. J. Welk (Ed.), *Physical activity assessments for health-related research* (pp. 125–141). Champaign, IL: Human Kinetics.

Authors' Note

Please address all correspondence concerning this article to Minsoo Kang, Department of Health and Human Performance, Middle Tennessee State University, P.O. Box 96, Murfreesboro, TN 37132.

E-mail: mkang@mtsu.edu