Estimating Characteristics of Energy Expenditure using Measurement Error Models

Nick Beyler* Sarah Nusser Wayne Fuller Alicia Carriquiry
Greg Welk

Abstract
One important aspect of physical activity research is the assessment of usual (i.e., long-term average) daily energy expenditure. Daily measurements of energy expenditure taken from a sample of individuals are prone to measurement errors and nuisance effects, which can lead to biased estimates of usual daily energy expenditure parameters. Fortunately, statistical models can be used to account and adjust for these errors in order to give more accurate estimates. In this paper we develop a method for estimating usual daily energy expenditure parameters from data collected using a self-report instrument and an objective monitoring device. Our method is an extension of existing methods that utilize measurement error models. We illustrate our method with preliminary data from the Physical Activity Measurement Survey (PAMS) collected using a SenseWear Pro armband monitor and a 24-hour physical activity recall.

Key Words: Physical activity, Energy expenditure, Measurement error model; PAMS

1. Introduction

Assessment of usual or habitual physical activity is important for studying relationships between physical activity and health and for determining appropriate physical activity guidelines to maintain good health (Shephard 2003). One component of this assessment involves estimation of usual daily energy expenditure (EE) parameters. EE is a measure of the energy cost associated with physical activity (Schutz et al. 2001). An individual's usual daily EE is his or her average daily EE over a long period of time, such as one year. From a statistical perspective, usual daily EE of individual i is

$$T_i = E\{T_{ij}|i\},$$

where T_{ij} is the actual daily EE of individual i on day j.

The instruments most commonly used to measure daily EE from individuals in the population are self-report instruments (Ainsworth 2009; Matthews 2002) and monitoring devices (Welk 2002; Moy et al. Submitted), both of which provide imperfect measurements of usual daily EE. An observed measurement of daily EE for individual i on day j, defined as Y_{ij}, will differ from the usual daily EE for individual i, T_i, because of nuisance effects (Matthews et al. 2001; Matthews et al. 2002) and measurements errors (Ainsworth 2009; Welk 2002). Nuisance effects, such as seasonality and day-of-week effect, exist because individuals vary their physical activity habits on a daily basis. Measurement errors from monitoring devices are due to the inability of monitors to accurately capture the full range of activities (Welk et al. 2004) and the imperfect conversion process of monitor data into EE estimates (Welk 2002). Measurement errors from self-report instruments are due to such factors as social desirability effects (Adams et al. 2005), difficulty in understanding

*Iowa State University, Ames, IA 50011

concepts of survey questions (Sallis and Saelens 2000), and cognitive limitations for recalling activity from the past (Matthews 2002). The difference between actual daily EE and usual daily EE may be defined as

$$D_{ij} = T_{ij} - T_i$$

for individual i on day j, and can be attributed to nuisance factors. For example, if individual i was more active than he or she usually is on day j, then $D_{ij} > 0$. The difference between measured and actual daily EE may be defined as

$$E_{ij} = Y_{ij} - T_{ij},$$

and can be attributed to measurement errors. For example, if individual i reports more activity than he or she actually did on day j using a self-report instrument, then $E_{ij} > 0$. The total difference between observed EE (Y_{ij}) and usual daily EE (T_i) is then

$$\begin{aligned} Y_{ij} - T_i &= T_{ij} - T_i + Y_{ij} - T_{ij} \\ &= D_{ij} + E_{ij}, \end{aligned}$$

for individual i on day j, which is the sum of the nuisance effect (D_{ij}) and the measurement error effect (E_{ij}). Failure to account for the measurement error and nuisance effects in daily EE measurements may lead to biased estimates of usual daily EE parameters. See, for example, Troiano et al. (2008) and Ferrari et al. (2007).

In this paper, we develop a method for estimating usual daily EE parameters that accounts for the measurement error and nuisance effects in observed EE data. In our method, parameters of usual daily EE are estimated from a sample of individuals in the population, where each individual provides replicate concurrent measurements of daily EE using a reference instrument, such as a multi-sensor monitor, and a self-report instrument, such as a 24-hour recall. Like other methods in the literature for estimating usual physical activity and dietary intake variables (Ferrari et al. 2007; Nusser et al. 1996; Kipnis et al. 2003), our method adjusts for the measurement error and nuisance effects associated with observed values of EE using measurement error models. Our method also includes a procedure for estimating usual daily EE parameters simultaneously for distinct groups in the population, which may be defined by gender, age, and race/ethnicity. This extension allows researchers to compare EE across groups that are of interest in physical activity assessment.

Our method consists of two general steps, which we briefly outline in this section. See Beyler (2010) for a more detailed description of these steps. The steps are used to estimate and remove measurement error and bias in the EE data before estimating usual daily EE parameters.

In the first step of our method, we transform the EE data to approximate normality and test for the presence of a variety of nuisance factors. In our analyses (Section 3), a log transformation gives approximately normal data, but in other cases, a power transformation or a more complex semiparametric transformation may be necessary to approximate normality. The transformation is important because the normality assumption is required to model the distribution of usual daily EE. We test for nuisance effects in the transformed data by fitting separate linear regression models to the EE measurements from the reference instrument and self-report instrument, which include nuisance effects parameters. Common nuisance effects to consider are day-of-week effect (e.g., weekday vs. weekend), time-in-sample effect (e.g., first vs. second replicate), and seasonality (e.g., summer vs. winter). If a nuisance effect is significant in the fitted linear regression models, the estimated effect is removed from the EE data and the remainder of the analyses are conducted with the adjusted EE data. If a nuisance effect is non-significant in the fitted models, the EE data are not adjusted for that effect. In out analyses (Section 3), no adjustments were made for nuisance effects because nuisance effect variables were non-significant in fitted regression models.

In the next step of our method, models are fit to the (possibly adjusted) normal-scale EE data to estimate sources of variation and bias in the data and to estimate parameters of usual daily EE. Assessment of usual daily EE in subpopulations (hereafter referred to as groups) is often of interest to public health researchers. In our method, groups can be defined by gender, age, race/ethnicity, or other factors with the goal of comparing model parameters for EE behaviors across these groups. After the groups are identified, a group-level measurement error model is fit to each group using method of moments. The same measurement error model is fit to each group so that parameter estimates can be compared across groups. A population-level model is then developed based on the group-level estimates so that the total number of model parameters may be reduced. If there is evidence that a group-level model parameter is similar across groups, the parameter may be pooled across the groups. If there is evidence of a systematic trend in a group-level parameter across groups, the trend can be accounted for with fewer parameters in the population-level model. Once the population-level model is specified, the model is fit to group-level moment estimators using estimated generalized least squares and estimated daily EE parameters are obtained.

The group-level and population-level models are developed in Section 2. We illustrate our method by estimating usual daily EE parameters from a preliminary sample of females in the Physical Activity Measurement Survey (PAMS) in Section 3. We give a discussion of the results in Section 4.

2. Models

The primary step in our method is parameter estimation for a group-level measurement error model and a population-level model. The group-level model is used to estimate daily EE parameters for each group. Groups may be defined by gender, age, race/ethnicity or any other factors of interest to the researcher. The population-level model is used to estimate daily EE parameters for the population by reducing the total number of model parameters. In this section, we present the group-level measurement error model and the population-level model. In what follows, assume that the EE data are in the normal scale (i.e., a transformation is applied to approximate normality).

2.1 Group-Level Model

Assume that G groups are considered for the analyses, and let g denote the gth group. Further assume that the EE measurements from group g and group g' are uncorrelated for $g \neq g'$. Let μ_g be the mean of daily EE in the normal scale for group g and let $\mu_g + t_{gi}$ be the mean daily EE for individual i in the normal scale, where $t_{gi} \sim N(0, \sigma^2_{tg})$. The distribution of mean daily EE in the normal scale is then given by $N(\mu_g, \sigma^2_{tg})$ for group g. On any given day j, individual i in group g

will have an actual daily EE value of t_{gij} in the normal scale. We assume that the daily deviations from the individual's mean daily EE are additive. Thus, our model for t_{gij} is

$$t_{gij} = \mu_g + t_{gi} + d_{gij},$$

where $d_{gij} \sim N(0, \sigma^2_{dg})$ is individual i's deviation from his or her mean daily EE on day j in the normal scale. On days where individual i is more active than usual, d_{gij} will be positive, and on days where individual i is less active than usual, d_{gij} will be negative. We assume that t_{gi} and d_{gij} are uncorrelated for all g, i, and j. That is, we assume that an individual's mean activity is unrelated to his or her within-individual variation in activity on a day-to-day basis.

Let x_{gij} be a measure of daily EE in the normal scale for individual i on day j in group g from an unbiased reference instrument, such as a multi-sensor monitoring device. We assume that the reference instrument gives an unbiased measurement of daily EE in the normal scale,

$$x_{gij} = \mu_g + t_{gi} + d_{gij} + u_{gij}, \tag{1}$$

where $u_{gij} \sim N(0, \sigma^2_{ug})$ is random measurement error for individual i on day j in group g. We assume that u_{gij} is uncorrelated with t_{gi} and d_{gij} for all g, i, and j, and hence, the variance of x_{gij} is the sum of three variance components,

$$\begin{aligned} V\{x_{gij}\} &= V\{\mu_g + t_{gi} + d_{gij} + u_{gij}\} \\ &= \sigma^2_{tg} + \sigma^2_{dg} + \sigma^2_{ug}. \end{aligned}$$

Let y_{gij} be a measurement of daily EE in the normal scale for individual i on day j in group g from a self-report instrument such as a 24-hour recall. We assume that the self-report measure y_{gij} is potentially biased for actual daily EE in the normal scale and represent y_{gij} as

$$y_{gij} = \mu_{yg} + \beta_{1g}(t_{gi} + d_{gij}) + r_{gi} + e_{gij}, \tag{2}$$

where μ_{yg} is the group mean of daily EE in the normal scale from the self-report instrument, β_{1g} is the slope that accounts for the systematic error in the relationship between self-report and actual daily EE in group g, $r_{gi} \sim N(0, \sigma^2_{rg})$ is a term that represents individual i's deviation from the group-level mean, and $e_{gij} \sim N(0, \sigma^2_{eg})$ is the remaining measurement error in the self-report for individual i on day j in group g. We assume that the model terms r_{gi} and e_{gij} are uncorrelated with each other, with t_{gi} and d_{gij}, and with u_{gij} from model (1) for all g, i, and j. Like model (1), model (2) assumes an additive linear relationship between measured EE and mean daily EE in the normal scale. Unlike model (1), model (2) includes a different overall mean, μ_{yg}, and a slope term, β_{1g}, to account for systematic error that may arise from self-reporting EE.

We use method of moments to derive estimators of the parameters for the group-level measurement error model given by (1) and (2). The estimators are given as weighted estimators, where w_{gi} is the weight for individual i in group g reflecting individual i's probability of selection into the sample. When weights are unavailable, w_{gi} is set to 1 for all g and i. Let the 8-dimensional parameter vector for group g be defined by

$$\boldsymbol{\theta}_g = (\mu_g, \mu_{yg}, \beta_{1g}, \sigma^2_{tg}, \sigma^2_{dg}, \sigma^2_{ug}, \sigma^2_{eg}, \sigma^2_{rg})'. \tag{3}$$

To compute estimators for $\boldsymbol{\theta}_g$, we consider summary statistics based on

$$\boldsymbol{Z}_{gi} = \begin{pmatrix} \bar{x}_{gi.} \\ \bar{y}_{gi.} \\ x_{gi1} - x_{gi2} \\ y_{gi1} - y_{gi2} \end{pmatrix}, \tag{4}$$

where

$$\bar{x}_{gi.} = \frac{x_{gi1} + x_{gi2}}{2}$$

and

$$\bar{y}_{gi.} = \frac{y_{gi1} + y_{gi2}}{2}.$$

We define $\boldsymbol{Z}_{gi}$ in this manner because $\boldsymbol{Z}_{gi}$ provides an algebraically simpler covariance matrix than the observed data vector $(x_{gi1}, x_{gi2}, y_{gi1}, y_{gi2})'$. Given the model assumptions, the expected value of $\boldsymbol{Z}_{gi}$ is

$$E\{\boldsymbol{Z}_{gi}\} = \begin{pmatrix} \mu_g \\ \mu_{yg} \\ 0 \\ 0 \end{pmatrix} \tag{5}$$

and the variance of $\boldsymbol{Z}_{gi}$ is

$$\begin{pmatrix} \sigma^2_{gt} + \frac{1}{2}\sigma^2_{gd} + \frac{1}{2}\sigma^2_{gu} & \beta_{1g}\sigma^2_{tg} + \frac{1}{2}\beta_{1g}\sigma^2_{dg} & 0 & 0 \\ & \beta^2_{1g}(\sigma^2_{tg} + \frac{1}{2}\sigma^2_{dg}) + \sigma^2_{rg} + \frac{1}{2}\sigma^2_{eg} & 0 & 0 \\ & & 2(\sigma^2_{dg} + \sigma^2_{ug}) & 2\beta_{1g}\sigma^2_{dg} \\ \textit{symmetric} & & & 2(\beta^2_{1g}\sigma^2_{dg} + \sigma^2_{eg}) \end{pmatrix}. \tag{6}$$

The sample mean of $\boldsymbol{Z}_{gi}$ is

$$\boldsymbol{m}_{1g} = \begin{pmatrix} m_{1g} \\ m_{2g} \\ 0 \\ 0 \end{pmatrix}, \tag{7}$$

where

$$\begin{aligned} m_{1g} &= \frac{\sum_{i=1}^{n_g} w_{gi} x_{gi.}}{\sum_{i=1}^{n_g} w_{gi}}, \\ m_{2g} &= \frac{\sum_{i=1}^{n_g} w_{gi} y_{gi.}}{\sum_{i=1}^{n_g} w_{gi}}, \end{aligned}$$

and n_g is the number of individuals in group g. The sample variance of $\boldsymbol{Z}_{gi}$ is

$$\boldsymbol{m}_{2g} = \frac{\sum_{i=1}^{n_g} w_{gi}(\boldsymbol{Z}_{gi} - \bar{\boldsymbol{Z}}_g)(\boldsymbol{Z}_{gi} - \bar{\boldsymbol{Z}}_g)'}{\sum_{i=1}^{n_g} w_{gi}},$$

where

$$\bar{\boldsymbol{Z}}_g = \frac{\sum_{i=1}^{n_g} w_{gi}\boldsymbol{Z}_{gi}}{\sum_{i=1}^{n} w_{gi}}$$

is the group sample mean of the $\boldsymbol{Z}_{gi}$. For deriving the method of moments estimating equations, we write

$$\boldsymbol{m}_{2g} = \begin{pmatrix} m_{11g} & m_{12g} & 0 & 0 \\ & m_{22g} & 0 & 0 \\ & & m_{33g} & m_{34g} \\ sym. & & & m_{44g} \end{pmatrix}, \tag{8}$$

where the sample moments m_{13g}, m_{14g}, m_{23g}, and m_{24g} are set to zero since their corresponding population moments in (6) are all zero.

The estimating equations are

$$\boldsymbol{m}_{1g} = E\{\boldsymbol{Z}_{gi}\}$$

and

$$\boldsymbol{m}_{2g} = V\{\boldsymbol{Z}_{gi}\},$$

where $\boldsymbol{m}_{1g}$ and $\boldsymbol{m}_{2g}$ are defined by (7) and (8), respectively, and $E\{\boldsymbol{Z}_{gi}\}$ and $V\{\boldsymbol{Z}_{gi}\}$ are defined by (5) and (6), respectively. There are eight model parameters and eight unique first and second moments in these equations, which allows for identification of each model parameter as a function of the sample moments. The method of moments estimators are given in Table 1. In what follows, we let

$$\hat{\boldsymbol{\theta}}_g = (\hat{\mu}_g, \hat{\mu}_{yg}, \hat{\beta}_{1g}, \hat{\sigma}^2_{tg}, \hat{\sigma}^2_{dg}, \hat{\sigma}^2_{ug}, \hat{\sigma}^2_{eg}, \hat{\sigma}^2_{rg})' \tag{9}$$

denote the method of moments estimator for the parameter vector $\boldsymbol{\theta}_g$ in (3).

Table 1: Method of moments estimators for group g

Parameter	Estimator
μ_g	$\hat{\mu_g} = m_{1g}$
μ_{yg}	$\hat{\mu}_{yg} = m_{2g}$
β_{1g}	$\hat{\beta}_{1g} = (m_{12g} - 0.25m_{34g})/(m_{11g} - 0.25m_{33g})$
σ^2_{tg}	$\hat{\sigma}^2_{tg} = m_{11g} - 0.25m_{33g}$
σ^2_{dg}	$\hat{\sigma}^2_{dg} = [m_{34g}(m_{11g} - 0.25m_{33g})]/[2(m_{12g} - 0.25m_{34g})]$
σ^2_{ug}	$\hat{\sigma}^2_{ug} = 0.5m_{33g} - [m_{34g}(m_{11g} - 0.25m_{33g})]/[2(m_{12g} - 0.25m_{34g})]$
σ^2_{eg}	$\hat{\sigma}^2_{eg} = 0.5m_{44g} - [m_{34g}(m_{12g} - 0.25m_{34g})]/[2(m_{11g} - 0.25m_{33g})]$
σ^2_{rg}	$\hat{\sigma}^2_{rg} = m_{22g} - 0.25m_{44g} - [(m_{12g} - 0.25m_{34g})^2]/[m_{11g} - 0.25m_{33g}]$

A Taylor series approximation is used to derive an estimated variance matrix for $\hat{\boldsymbol{\theta}}_g$. The approximation is given by

$$\hat{V}\{\hat{\boldsymbol{\theta}}_g\} = \hat{\boldsymbol{D}}_g \hat{V}\{\boldsymbol{m}_g\}\hat{\boldsymbol{D}}'_g, \tag{10}$$

where $\hat{\boldsymbol{D}}_g$ is a matrix of derivatives for the method of moments estimators evaluated at the method of moments estimates and $\hat{V}\{\boldsymbol{m}_g\}$ is an estimated variance of the sample moments

$$\boldsymbol{m}_g = (m_{1g}, m_{2g}, m_{11g}, m_{12g}, m_{22g}, m_{33g}, m_{34g}, m_{44g})'. \tag{11}$$

To derive the matrix of derivatives, let m_{gk} denote the kth element in $\boldsymbol{m}_g$ for $k = 1, \ldots, 8$ and let $b_l(\boldsymbol{m}_g)$ be a function of $\boldsymbol{m}_g$ that represents the lth method of moments estimator in Table 1 for $l = 1, \ldots, 8$. Then, define $\hat{\boldsymbol{D}}_g$ to be an 8 x 8 matrix of derivatives for the sample moments, where element lk in $\hat{\boldsymbol{D}}_g$ is

$$D_{glk} = \frac{\partial b_l(\boldsymbol{m}_g)}{\partial m_{gk}}$$

for $l = 1, \ldots, 8$ and $k = 1, \ldots, 8$. The variance of $\boldsymbol{m}_g$ can be estimated using a Horvitz-Thompson variance to account for the sample design. The Horvitz-Thompson variance estimator is

$$\hat{V}\{\boldsymbol{m}_g\} = \sum_{i=1}^{n_g} \sum_{k=1}^{n_g} \pi_{ik}^{-1}(\pi_{ik} - \pi_i \pi_k) w_{gi} \boldsymbol{s}_{gi} w_{gk} \boldsymbol{s}'_{gk},$$

where π_i is the first order inclusion probability of individual i into the sample, π_{ik} is the second order inclusion probability of individuals i and k into the sample, w_{gi} is the survey weight for individual i in group g, and

$$\boldsymbol{s}_{gi} = \begin{pmatrix} \bar{x}_{gi.} \\ \bar{y}_{gi.} \\ (\bar{x}_{gi.} - m_{1g})^2 \\ (\bar{x}_{gi.} - m_{1g})(\bar{y}_{gi.} - m_{2g}) \\ (\bar{y}_{gi.} - m_{2g})^2 \\ (x_{gi1} - x_{gi2})^2 \\ (x_{gi1} - x_{gi2})(y_{gi1} - y_{gi2}) \\ (y_{gi1} - y_{gi2})^2 \end{pmatrix} \tag{12}$$

is the vector of summary statistics for individual i.

2.2 Population-Level Model

In the previous section we developed estimators which can be used to estimate the group-level model parameters, including the group mean (μ_g) and variance (σ^2_{tg}) of daily EE in the normal scale. Although it is of interest to estimate separate parameters for each of the G groups, it is possible that the group-level parameters can be modeled across the groups to form a population-level model with a reduced number of parameters. In this section, we outline a procedure for developing a population-level model from the group-level model parameters. In this section we give the general form of the model and an estimator for the model parameter vector. In Section 3 we illustrate how the model can be formulated.

The population-level model is defined by a set of functions that model the group-level parameters in $\boldsymbol{\theta}_g$ given by (3) as functions of a new set of parameters defined for the population. The set of functions and population-level model parameters are formulated based on an analysis of the group-level parameter estimates. The general form of the population-level model is

$$\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\lambda} + \boldsymbol{e}. \tag{13}$$

In the model,

$$\boldsymbol{y} = (\hat{\boldsymbol{\theta}}'_1, \ldots, \hat{\boldsymbol{\theta}}'_G)'$$

is the $8G$-dimensional vector of the estimated group-level model parameters, where $\hat{\boldsymbol{\theta}}_g$ is given by (9) for group g, $g = 1, \ldots, G$. $\boldsymbol{\lambda}$ is the q-dimensional vector of parameters for model (13), where $q < 8G$ so that the total number of parameters from the group-level models is smaller for the population-level model. $\boldsymbol{Z}$ is a ($8G$ x q) design matrix for the model representing coefficients that define the set of functions that relate the $8G$ group-level estimated parameters to linear functions of the q population-level parameters. The variance of the vector of error terms, $\boldsymbol{e} \sim (\boldsymbol{0}, \boldsymbol{V})$, is estimated by

$$\hat{\boldsymbol{V}} = blockdiag(\hat{V}\{\hat{\boldsymbol{\theta}}_1\}, \ldots, \hat{V}\{\hat{\boldsymbol{\theta}}_G\}), \tag{14}$$

where $\hat{V}\{\hat{\boldsymbol{\theta}}_g\}$ is given by (10) for $g = 1, \ldots, G$. The estimated variance (14) is appropriate under the assumption that the EE measurements are uncorrelated across groups. With an estimated variance $\hat{\boldsymbol{V}}$, the population-level model can be estimated using estimated generalized least squares (EGLS). The EGLS estimator of $\boldsymbol{\lambda}$ is

$$\hat{\boldsymbol{\lambda}} = (\boldsymbol{Z}'\hat{\boldsymbol{V}}^{-1}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\hat{\boldsymbol{V}}^{-1}\boldsymbol{y}, \tag{15}$$

and an estimated variance of the estimator is

$$\hat{V}\{\hat{\boldsymbol{\lambda}}\} = (\boldsymbol{Z}'\hat{\boldsymbol{V}}^{-1}\boldsymbol{Z})^{-1}. \tag{16}$$

Using the population-level model, we can estimate model parameters for each group. Let $\hat{\mu}_g$ and $\hat{\sigma}^2_{tg}$ denote the estimated mean daily EE and estimated variance of daily EE in the normal scale for group g, $g = 1, \ldots, G$, from the population-level model. The estimated distribution of mean daily EE in the normal scale for group g is then $N(\hat{\mu}_g, \hat{\sigma}^2_{tg})$. Estimates of other model parameters can also be obtained using the population-level model, including estimates of the slope parameters relating actual daily EE to self-reported EE ($\hat{\beta}_{1g}$), estimates of the group means of self-reported EE ($\hat{\mu}_{yg}$), and estimates of the variance components that account for day-to-day variation in daily EE ($\hat{\sigma}^2_{dg}$), measurement error variation in the reference instrument ($\hat{\sigma}^2_{ug}$) and self-report instrument ($\hat{\sigma}^2_{eg}$), and random variation due to self-reporting ($\hat{\sigma}^2_{rg}$).

3. Application to PAMS Data

In this section, we use the method described in Sections 1 and 2 to estimate usual daily EE parameters using preliminary EE data from the Physical Activity Measurement Survey (PAMS). The preliminary data come from a sample of 171 females selected into the PAMS sample who provided concurrent replicate measurements of daily EE using the SenseWear armband monitor and a 24-hour physical activity recall. See Beyler (2010) for a description of the PAMS survey design and descriptive statistics of the preliminary EE data.

To implement our method, we first transform the daily EE data to approximate normality. Because log transformations are often used for analyses in physical activity research (Ferrari et al. 2007), we consider the log transformation to approximate normality for the PAMS EE data. Let $x_{ij} = log(X_{ij})$ be daily EE from the monitor and let $y_{ij} = log(Y_{ij})$ be daily EE from the 24PAR in the log scale for individual i on day j. Shapiro-Wilk test statistics are computed for the set of x_{ij} values and set of y_{ij} values from the sample using SAS statistical software (SAS Institute 2009). The p-values for the test statistics are 0.25 and 0.21 for the set of x_{ij} and y_{ij} values, respectively. Therefore, the log transformed values are used for model fitting.

Next, we check for nuisance effects in the log-transformed daily EE data by fitting linear regression models containing covariates for day-of-week effect, time-in-sample effect, and demographic variables. We include variables for day-of-week effect and time-in-sample effect in the models because we suspect that an individual may have different EE values depending on the day of the week (e.g., weekday vs. weekend) and depending on whether the value is the first or second observation for a respondent (e.g., replicate 1 vs. 2). We include demographic variables for age, race/ethnicity, education, and smoking status in the models because we suspect that EE levels may vary by these factors. The nuisance effects variables were non-significant in the fitted models, and as a result, no additional adjustments were made to the log-transformed daily EE data.

In preliminary analyses, daily EE measurements from the monitor were shown to vary according to age (results not shown). Based on these results, we define age groups for the group-level measurement error models. To form groups, we divide the sample into four age groups of approximately equal size (Table 2). For the remainder of the presentation, we will denote the age groups as groups 1 - 4, where 1 is the youngest age group and 4 is the oldest age group.

Table 2: Age groups

Age Group g	Age Range	Average Age	Sample Size
1	23 - 42	34.3	44
2	43 - 52	48.6	40
3	53 - 59	55.6	43
4	60 - 70	64.6	44

Once the groups have been determined, the next step in our method is to estimate the group-level model parameters. The measurement error model given by equations (1) and (2) is fit to each of the four age groups using method of moments as described in Section 2. The method of moments estimators are given in Table 1 in Section 2. Standard errors for the parameter estimates are computed using the Taylor series variance estimator given by (10), $\hat{V}\{\hat{\boldsymbol{\theta}}_g\} = \boldsymbol{D}_g \hat{V}\{\boldsymbol{m}_g\} \boldsymbol{D}'_g$. Due to the small number of individuals in each of the 4 age groups, we ignore the survey design in computing the estimated variance of the sample moments, $\hat{V}\{\boldsymbol{m}_g\}$, and instead use the estimated variance for a simple random sample (ignoring the finite population correction) defined by

$$\hat{V}\{\boldsymbol{m}_g\} = n_g^{-1}(n_g - 1)^{-1} \sum_{i=1}^{n_g} (\boldsymbol{s}_{gi} - \bar{\boldsymbol{s}}_{g.})(\boldsymbol{s}_{gi} - \bar{\boldsymbol{s}}_{g.})',$$

where $\boldsymbol{s}_{gi}$ is given by (12) and $\bar{\boldsymbol{s}}_{g.}$ is the mean of the $\boldsymbol{s}_{gi}$ in group g. The parameter estimates and standard errors from the measurement error models for each group are given in Tables 3 and 4.

Table 3 contains the estimated group-level measurement error model parameters for the mean of daily EE (μ_g), the mean of reported daily EE (μ_{yg}) and the slope relating mean daily EE to reported daily EE (β_{1g}). The estimated means of daily EE decrease by age group, suggesting that older females tend to have lower levels of mean daily EE compared to younger females. The estimated slope parameters also decrease by age group, suggesting that the relationship between average levels of mean daily EE and reported daily EE may be a function of age. The estimated means of reported daily EE are larger than the estimated means of daily EE, suggesting over-reporting in daily EE for all age groups. Unlike the daily EE means, the reported daily EE means do not show much of a trend across age groups. Given these results, we model the decreasing trends in the estimated daily EE means and the estimated slope parameters in the population-level model and estimate a common mean for reported daily EE.

Table 3: Estimated measurement error model parameters (and standard errors) for the mean of daily EE (μ_g), the mean of reported daily EE (μ_{yg}), and the slope for population-level reporting bias (β_{1g})

Parameter	Group 1	Group 2	Group 3	Group 4
μ_g	7.8421 (.0240)	7.8104 (.0230)	7.7595 (.0283)	7.7182 (.0241)
μ_{yg}	8.0616 (.0435)	8.0300 (.0326)	8.0656 (.0398)	8.0318 (.0315)
β_{1g}	1.2970 (.2103)	0.9226 (.1984)	0.8433 (.2081)	0.6982 (.1145)

Table 4 contains the estimated group-level measurement error model parameters for the variance components from models (1) and (2). No systematic trends in the components are discernible. In preliminary analysis of possible models for the variance components, the largest differences between age groups were non-significant (results not shown). As more data become available, evidence of relationships or differences in the variance components across age groups may surface. For this analyses, we assume constant variance components across age groups in the population-level model.

Table 4: Estimated variance components (and standard errors) from the measurement error model for mean daily EE (σ^2_{tg}), within-individual EE variation in daily EE (σ^2_{dg}), measurement error variation from the monitor (σ^2_{ug}) and the 24PAR (σ^2_{eg}), and reporting-bias variation from the recall (σ^2_{rg})

Parameter	Group 1	Group 2	Group 3	Group 4
σ^2_{tg}	0.0211 (.0059)	0.0170 (.0048)	0.0296 (.0077)	0.0235 (.0062)
σ^2_{dg}	0.0089 (.0027)	0.0044 (.0022)	0.0025 (.0015)	0.0065 (.0028)
σ^2_{ug}	0.0056 (.0028)	0.0047 (.0027)	0.0067 (.0028)	0.0044 (.0034)
σ^2_{eg}	0.0042 (.0041)	0.0066 (.0022)	0.0064 (.0024)	0.0079 (.0028)
σ^2_{rg}	0.0363 (.0088)	0.0221 (.0049)	0.0416 (.0095)	0.0259 (.0085)

Given the results from the fitted group-level models (Table 3 and 4), we develop a population-level model for daily EE. We model the daily EE mean for age group g as

$$\mu_g = \mu_0 + \theta A_g, \tag{17}$$

where μ_0 is a baseline parameter for the daily EE mean in the population, A_g is the mean age of age group g minus the overall mean age for the sample, and θ is a parameter to estimate the linear trend in the daily EE mean. We model the bias slope parameters as a function of mean age,

$$\beta_{1g} = \beta_1 + \beta_3 A_g, \tag{18}$$

where β_1 is the baseline slope for the population and β_3 accounts for the linear trend in the slopes across age groups. We model the group means of reported EE as

$$\mu_{yg} = \mu_y + \beta_{1g}(\mu_g - \mu_0),$$

where μ_y is the overall mean of reported EE and $\beta_{1g}(\mu_g - \mu_0)$ accounts for the deviation in the group-level reported EE mean from the overall mean. Given models

(17) and (18), the model for the mean of reported EE can be written as

$$\mu_{yg} = \mu_y + (\beta_1 + \beta_3 A_g)\theta A_g. \tag{19}$$

The group-level variance components are related to population-level variance components through the system of equations

$$\begin{aligned} \sigma^2_{tg} &= \sigma^2_t \\ \sigma^2_{dg} &= \sigma^2_d \\ \sigma^2_{ug} &= \sigma^2_u \\ \sigma^2_{eg} &= \sigma^2_e \\ \sigma^2_{rg} &= \sigma^2_r, \end{aligned} \tag{20}$$

for $g = 1, \ldots, 4$.

The population-level model is given by (13) in Section 2. The 32-dimensional vector $\boldsymbol{y}$ is

$$\boldsymbol{y} = (\hat{\boldsymbol{\theta}}'_1, \hat{\boldsymbol{\theta}}'_2, \hat{\boldsymbol{\theta}}'_3, \hat{\boldsymbol{\theta}}'_4)',$$

where

$$\hat{\boldsymbol{\theta}}_g = (\hat{\mu}_g, \hat{\mu}_{yg}, \hat{\beta}_{1g}, \hat{\sigma}^2_{tg}, \hat{\sigma}^2_{dg}, \hat{\sigma}^2_{ug}, \hat{\sigma}^2_{eg}, \hat{\sigma}^2_{rg})'$$

is the estimated 8-dimensional vector of group-level model parameters for age group g, $g = 1, \ldots, 4$. The 10-dimensional vector of population-level model parameters is

$$\boldsymbol{\lambda} = (\mu_0, \mu_y, \theta, \beta_1, \beta_3, \sigma^2_t, \sigma^2_d, \sigma^2_u, \sigma^2_e, \sigma^2_r)',$$

where the parameters are defined in equations (17), (18), (19), and (20). The 32 x 10 design matrix $\boldsymbol{Z}$ is given by

$$\boldsymbol{Z} = \begin{pmatrix} \boldsymbol{Z}_1 \\ \boldsymbol{Z}_2 \\ \boldsymbol{Z}_3 \\ \boldsymbol{Z}_4 \end{pmatrix}$$

where

$$\boldsymbol{Z}_g = \begin{pmatrix} 1 & 0 & A_g & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & \hat{\beta}_{1g} A_g & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & A_g & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

for age group g. Note that the design matrix uses the linear approximation

$$\hat{\mu}_{yg} = \mu_y + \hat{\beta}_{1g}\theta A_g$$

for the nonlinear model

$$\hat{\mu}_{yg} = \mu_y + (\beta_1 + \beta_3 A_g)\theta A_g.$$

The population-level model is fit using the EGLS estimator (15), and the variance is estimated by (16). The parameter estimates and standard errors (computed from the EGLS variance) are given in Table 5. Each of the model parameters is significant at the 0.05 level. There is evidence of a linear trend across age groups in the daily EE mean (represented by θ), and evidence of a linear trend across age groups in the slope parameter (represented by β_3). The estimated mean of daily EE (μ_0) appears to be smaller than the estimated mean of reported daily EE (μ_y), indicating over-reporting bias in daily EE from the 24PAR. The estimated variance for individual reporting effects (σ_r^2) is large relative to the other estimated variances components. The estimated inter-individual variance in usual daily EE (σ_t^2) is about 4 times larger than the estimated within-individual variance in daily EE (σ_d^2).

Table 5: Parameter estimates (standard errors) for the population-level model

Parameter	Est	(SE)
μ_0	7.7940	(.0099)
μ_y	8.0564	(.0145)
100θ	-0.2409	(.0838)
β_1	0.9950	(.0689)
$100\beta_3$	-1.7035	(.4517)
$100\sigma_t^2$	1.9868	(.2442)
$100\sigma_d^2$	0.4949	(.0892)
$100\sigma_u^2$	0.5186	(.1125)
$100\sigma_e^2$	0.6175	(.1172)
$100\sigma_r^2$	2.0934	(.2999)

The parameter estimates from the population-model can be used to estimate normal-scale mean daily EE values in each of the 4 age groups. The estimated means are computed from equation (17) as $\hat{\mu}_g = \hat{\mu}_0 + \hat{\theta} A_g$, where $\hat{\mu}_0$ and $\hat{\theta}$ are given in Table 5. Standard errors for the estimated means are given by

$$se(\hat{\mu}_g) \quad = \quad \sqrt{\boldsymbol{c}_g' \hat{V}\{\hat{\boldsymbol{\lambda}}\} \boldsymbol{c}_g},$$

where

$$\boldsymbol{c}_g' = (1, 0, A_g, 0, 0, 0, 0, 0, 0, 0)$$

and $\hat{V}\{\hat{\boldsymbol{\lambda}}\}$ is the estimated variance matrix for the population-level model. The estimates and standard errors are given in Table 6. The parameter estimates from the population-level model can also be used to estimate the slope parameters for each of the age groups based on equation (18). The estimates (and standard errors) for the slope parameters are given in Table 6. Note that the estimated means and slope parameters in Table 6 are similar to the estimated means and slope parameters in Table 2 for the fitted group-level models.

Figure 1 illustrates the relationships between mean daily EE and reported daily EE in the youngest and oldest age groups. In the youngest age group (group 1), females with higher levels of usual daily EE tend to have a greater discrepancy between their reported and mean daily EE, while in the oldest age group (group 4), females with higher levels of mean daily EE tend to have a smaller discrepancy between their reported and mean daily EE.

Table 6: Parameter estimates (standard errors) for the daily EE group means (μ_g) and slope parameters (β_{1g}) based on the fitted population-level model

Parameter	Est	(SE)
μ_1	7.8337	(.0153)
μ_2	7.7992	(.0097)
μ_3	7.7824	(.0114)
μ_4	7.7607	(.0167)
β_{11}	1.2760	(.1169)
β_{12}	1.0323	(.0728)
β_{13}	0.9134	(.0652)
β_{14}	0.7599	(.0764)

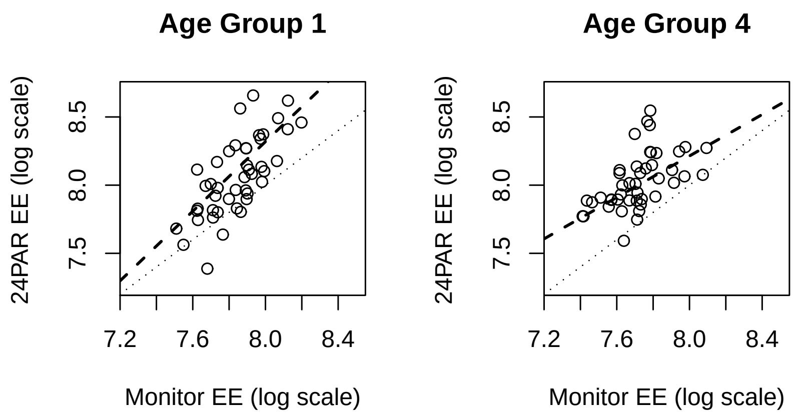

Figure 1: Estimated lines relating mean daily EE and reported daily EE for age groups 1 and 4 (points are the individual means of measured EE in the log scale; dashed lines are the estimated lines and dotted lines are the identity lines)

4. Discussion

In this paper, we have presented a method for estimating usual daily EE parameters, where daily EE measurements are adjusted for measurement error and nuisance effects using measurement error models. Our method is an extension of existing methods proposed in the literature for estimating usual physical activity parameters (Ferrari et al. 2007) and usual intake parameters (Nusser et al. 1996; Kipnis et al. 2003). A useful feature of our analysis is estimation of daily EE parameters for groups of the population. To implement our method, multiple concurrent measurements of daily EE must be available from an unbiased reference instrument, such as a multi-sensor monitoring device, and a self-report instrument, such as a 24-hour recall. The reference instrument is assumed to give unbiased measurements of usual daily EE for model identification purposes.

A number of interesting points were identified by the analysis of the PAMS data in Section 3. We estimated that a significant amount of the variation in daily EE measured from the 24PAR is due to individual-level reporting biases. Individuals tend to misreport on their daily EE from the previous day, which could be due to social desirability effects (Adams et al. 2005) or cognitive limitations associated with recalling activity from the past (Matthews 2002). Researchers should use caution when making inferences on self-reported EE data because of the potential for bias and excess variation in the data.

The results from the female PAMS sample also suggest that the within-individual variation in daily EE is small relative to the inter-individual variation in usual daily EE. In Table 5, the estimated usual daily EE variance in the normal scale (σ_t^2) is about 4 times larger than the estimated within-individual variance of daily EE (σ_d^2). This result is contrary to results from the dietary intake literature, which indicate that there is much more within-individual variation in dietary intake than there is inter-individual variation (Nusser et al. 1996; Carriquiry 2003).

In our analyses, there was evidence of a decrease in mean usual daily EE as age increases. The youngest age group (age 21 - 42) had the largest estimated mean of usual daily EE, while the oldest age group (age 60 - 70) had the smallest estimated mean of usual daily EE. Similar results are given in Ferrari et al. (2007), which show lower levels of estimated EE in older age groups relative to younger age groups. The estimated slope parameters, which compare usual daily EE to reported daily EE in the groups, also decreased with age. The more active females in the youngest age group tend to have larger discrepancies between their usual daily EE and reported daily EE, while the more active females in the oldest age group tend to have smaller discrepancies between their usual daily EE and reported daily EE (Figure 1).

A long-term goal of our research is to use estimated usual daily EE distributions to estimate usual daily EE parameters in the original scale and to infer about EE behaviors of individuals in the population. Some of this work has been explored in this paper and in Beyler (2010). Future work should involve a more thorough development of the methodology we have considered in this paper using EE data from a larger sample of the population.

References

[1] Adams, S.A., Matthews, C.E., Ebbeling, C.B., Moore, C.G., Cunningham, J.E., Fulton, J., and Hebert, J.R. (2005). The effect of social desirability and social approval on self-reports of physical activity. *American Journal of Epidemiology*, **161**, 389-398.

[2] Ainsworth, B.E. (2009). How do I measure physical activity in my patients? Questionnaires and objective methods. *British Journal of Sports Medicine*, **43**, 6-9.

[3] Beyler, N.K. (2010). Statistical methods for analyzing physical activity data. Unpublished PhD dissertation. Iowa State University, Ames, IA.

[4] Carriquiry, A.L. (2003). Estimation of usual intake distributions of nutrients and foods. *Journal of Nutrition*, **133**, 601S-608S.

[5] Ferrari, P., Friedenreich, C., and Matthews, C.E. (2007). The role of measurement error in estimating levels of physical activity. *American Journal of Epidemiology*, **166**, 832-840.

[6] Kipnis, V., Subar, A.F., Midthune, D., Freedman, L.S., Ballard-Barbash, R., Troiano, R.P., Bingham, S., Schoeller, D.A., Schatzkin, A., and Carroll, R.J. (2003). Structure of dietary measurement error: results of the OPEN Biomarker Study. *American Journal of Epidemiology*, **158**, 14-21.

[7] Matthews, C.E. (2002). Use of self-report instruments to assess physical activity. In: Welk, G.J. ed. *Physical Activity Assessments for Health Related Research*. Champaign: Human Kinetics.

[8] Matthews, C.E., Ainsworth, B.E., Thompson, R.W., and Bassett, D.R. (2002). Sources of variance in daily physical activity levels as measured by an accelerometer. *Medicine and Science in Sports and Exercise*, **34**, 1376-1381.

[9] Matthews, C.E., Hebert, J.R., Freedson, P.S., Stanek, E.J., Merriam, P.A., Ebbeling, C.B., and Ockene, I.S. (2001). Sources of variance in daily physical activity levels in the Seasonal Variation of Blood Cholesterol Study. *American Journal of Epidemiology*, **153**, 987-995.

[10] Moy, K.L., Calabro, M.A., McClain, J., and Welk, G.J. Comparison of multichannel pattern recognition monitors for estimating energy expenditure – a brief review. Submitted.

[11] Nusser, S.M., Carriquiry, A.L., Dodd, K.W., and Fuller, W.A. (1996). A semiparametric transformation approach to estimating usual daily intake distributions. *Journal of the American Statistical Association*, **91**, 1440-9.

[12] SAS Institute, Inc. (2009). SAS Online Doc, version 9.2. Cary, NC: SAS Institute, Inc.

[13] Sallis, J.F. and Saelens, B.E. (2000). Assessment of physical activity by self-report: status, limitations and future directions. *Research Quarterly for Exercise and Sport*, **71**, 1-14.

[14] Schutz, Y., Weinsier, R.L., and Hunter, G.R. (2001). Assessment of free-living physical activity in humans: an overview of currently available and proposed new measures. *Obesity Research*, **9**, 368-379.

[15] Shephard, R.J. (2003). Limit to the measurement of habitual physical activity by questionnaires. *British Journal of Sports Medicine*, **37**, 197-206.

[16] Troiano, R.P., Berrigan, D., Dodd, K.W., Masse, L.C., Tilert, T., and McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise*, **40**, 181-188.

[17] Welk, G.J. (2002). Use of accelerometry-based activity monitors to assess physical activity. In: Welk, G.J. ed. *Physical Activity Assessments for Health Related Research*. Champaign: Human Kinetics.

[18] Welk G.J., Schaben, J.A., and Morrow, J.R. (2004). Reliability of accelerometry-based activity monitors: a generalizability study. *Medicine and Science in Sports and Exercise*, **36**, 1637-1645.