Original Article
EPIDEMIOLOGY/GENETICS

Obesity

Illustration of Measurement Error Models for Reducing Bias in Nutrition and Obesity Research Using 2-D Body Composition Data

Anarina L. Murillo [1], *Olivia Affuso* [2,3,4], *Courtney M. Peterson* [2,5], *Peng Li* [1,6], *Howard W. Wiener*[3], *Carmen D. Tekwe* [7], *and David B. Allison* [8]

Objective: This study aimed to illustrate the use and value of measurement error models for reducing bias when evaluating associations between body fat and having type 2 diabetes (T2D) or being physically active.
Methods: Logistic regression models were used to evaluate T2D and physical activity among adults aged 19 to 80 years from the Photobody Study (n = 558). Self-reported T2D and physical activity were categorized as "yes" or "no." Body fat measured by two-dimensional photographs was adjusted for bias using dual-energy x-ray absorptiometry scans as a reference. Three approaches were applied: regression calibration (RC), simulation extrapolation (SIMEX), and multiple imputation (MI).
Results: Unadjusted two-dimensional measures of body fat had upward biases of 30% and 233% for physical activity and T2D, respectively. For the physical activity model, RC-adjusted values had a 13% upward bias, whereas MI and SIMEX decreased the bias to 9% and 91%, respectively. For the T2D model, MI reduced the bias to 0%, whereas RC and SIMEX increased the upward bias to > 300%.
Conclusions: Of three statistical approaches to reducing bias due to measurement errors, MI performed best in comparison to RC and SIMEX. Measurement error methods can improve the reliability of analyses that look for relations between body fat measures and health outcomes.

Obesity (2019) **27**, 489-495. doi:10.1002/oby.22387

Introduction

Measurement errors can manifest in health care research, particularly in obesity and nutrition studies in which self-reported measures are commonly used. It has been shown that self-reported measures, such as of dietary intake (1,2), physical activity levels, smoking behavior (3), and alcohol intake (4), are all prone to measurement error. These errors can arise from multiple sources and often lead to biased statistical inference and incorrect conclusions. In nutritional epidemiology, measurement errors have led to statistical bias when evaluating the relationship between self-reported energy intake assessed by the use of food frequency questionnaires (FFQs) (5-7) and chronic disease outcomes. Measurement errors can lead to biased estimates of the effects of error-prone measures on the outcomes of interest, loss of statistical power for detecting health outcomes due to potential excess variability, and an obscuring of the true features of the data (e.g., linear and nonlinear trends, associations between data variables) (8,9).

[1] Department of Biostatistics, University of Alabama at Birmingham, Birmingham, Alabama, USA. [2] Nutrition Obesity Research Center, University of Alabama at Birmingham, Birmingham, Alabama, USA [3] Department of Epidemiology, University of Alabama at Birmingham, Birmingham, Alabama, USA [4] Center for Exercise Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA [5] Department of Nutrition Science, University of Alabama at Birmingham, Birmingham, Alabama, USA [6] School of Nursing, University of Alabama at Birmingham, Birmingham, Alabama, USA [7] Department of Epidemiology and Biostatistics, Texas A&M University, College Station, Texas, USA [8] Department of Epidemiology and Biostatistics, School of Public Health, Indiana University-Bloomington, Bloomington, Indiana, USA. Correspondence: David B. Allison (allison@iu.edu)

Funding agencies: This research was in part funded by the National Heart, Lung, and Blood Institute (NHLBI) and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) of the National Institutes of Health (NIH) under Award Numbers R01HL107916 and P30DK056336. ALM acknowledges funding from NIDDK grant number T32DK062710 and NHLBI grant number T32HL072757. DBA acknowledges NIDDK grant number R25DK099080 and NHLBI grant number R25HL124208. CMP was supported by KL2TR001419 from the National Center for Advancing Translational Sciences (NCATS). CDT's research was supported by National Cancer Institute Supplemental Award Number U01-CA057030-29S2. The content is the sole responsibility of the authors and does not necessarily represent the official views of the NIH or any other organization.

Disclosure: The authors declared no conflicts of interest.

Author contributions: DBA and OA designed the research study; OA and HWW were responsible for data collection and data checking; DBA and ALM designed the statistical analyses; OA, CMP, PL, and CDT provided feedback and suggestions on statistical analyses; ALM conducted the statistical analyses and wrote the first draft of the manuscript; and PL and HWW validated the statistical analyses. All authors edited and contributed to writing the final draft and approved the final submitted manuscript. DBA, OA, and ALM had primary responsibility for the final content of the submitted manuscript.
Additional Supporting Information may be found in the online version of this article.

Received: 6 May 2018; **Accepted:** 1 November 2018; **Published online 22 January 2019. doi:10.1002/oby.22387**

Measurement error (systematic and random error) in health research can arise from multiple sources (e.g., heart rate may be prone to within-individual variability in repeated measures because of the instrument or physiology of an individual), and measurement error can manifest in different patterns (e.g., consistent underreporting or overreporting of self-reported variables). In nutritional epidemiology, for example, error-prone measurements of energy intake (such as FFQ data, which may be subject to inaccurate responses due to participants' inability to accurately recall food consumption) lead to biased estimates of the effects of diet on health outcomes (7).

Several statistical methods are available to reduce bias due to measurement error, including the classic regression calibration (RC) approach (10), multiple imputation (MI) (5,11,12), the maximum likelihood method (13), simulation extrapolation (SIMEX) (14,15), and other methods (16-18). Spiegelman et al. showed improvements in statistical inference for FFQ data following RC (10,19). Prentice et al. (6,20) showed that unadjusted parameters of energy intake obtained from FFQs were not significantly associated with cancer; however, after bias adjustment of measurement error using calibration methods, energy and protein density were, indeed, positively associated with cancer incidence (20). In this example, adjusting for measurement error improved the estimation of the association between diet and cancer, which further highlights the value of implementing these methods to improve estimation and reduce bias in statistical analyses. Similarly, calibrated or measurement error bias-adjusted energy consumption was positively associated with coronary artery disease risk (6), whereas unadjusted energy consumption was not.

Anthropometric measures, such as BMI (21,22) and body fat percentage measured by three-dimensional photonic scans or dual-energy x-ray absorptiometry (DXA) scans (23,24), are also prone to inherent or unavoidable measurement error. While many body composition assessment methods are available, such as densitometry methods (e.g., air displacement, underwater weighing) and bioimpedance analysis, DXA scans are often the preferred method for estimating body fat. Furthermore, Garlie et al. showed that body fat percentage as estimated by three-dimensional photonic scan yielded 4.69% to 5.99% error relative to DXA (25). Therefore, measurement error remains an issue in body composition data, and to our knowledge, the available statistical approaches for measurement error bias adjustment have not yet been commonly adopted in body fat measures.

The purpose of this study was to illustrate the use and value of these methods to reduce potential biases due to measurement error when assessing the effects of body fat on two health outcomes: (1) the probability of having type 2 diabetes (T2D) and (2) the probability of being physically active. In this study, we analyzed a unique data set with body fat percentage estimated in two ways for all participants. First, body fat percentage was estimated by use of a novel two-dimensional photographic-based method (BF_{Photo}), which was developed by Affuso et al. (26-28). This photographic method has been shown to provide a valid estimate of body fat percentage compared with DXA in adults. However, this novel method may involve some inherent and unavoidable measurement error due to potentially lost information about muscle mass, bone mass, and fat mass, which may be lost when extracting body volume from two-dimensional photographs and therefore may lead to biased estimates of body fat percentage. Second, body fat percentage was estimated by DXA (BF_{DXA}). The two-dimensional photographic method has several advantages over DXA, such as its portability, low cost, convenience, and time efficiency. While BF_{Photo} estimates have been shown to be strongly correlated to BF_{DXA} (26,28), here we assume that BF_{DXA} is the reference method, or is error free, while BF_{Photo} is an error-prone body fat measure. In this work, three measurement error bias-adjustment techniques were applied to improve the predictions of health outcomes as estimated by BF_{Photo} by improving model parameters enough to yield results similar to those of the reference method, BF_{DXA}.

Methods

Subjects

Adults aged 19 to 80 years living in Birmingham, Alabama, were recruited between November 2012 and September 2015 as part of the Photobody Study described in prior publications (26-28), and a subset of the data is used here. This subset included non-Hispanic black participants or non-Hispanic white participants only. Participants who reported a "moderate" level of regular physical activity were excluded to allow for greater contrast between the individuals with low and high activity. Participants were recruited through advertisements placed in local newspapers and newsletters, flyers were placed throughout the community (e.g., at college campuses and other businesses), individuals were directly approached at community events (e.g., at health fairs and other local events), and enrolled participants communicated to others by word of mouth. Individuals meeting the following criteria were enrolled in the study: (1) weight less than 450 lb (weight limit of DXA equipment), (2) absence of conditions that would prevent participants from lying down for DXA scans or standing for taking photographs, (3) presence of health conditions that may alter body composition (e.g., cancer, cachexia, rheumatoid arthritis), (4) no missing body parts (except a finger or toe), and (5) not pregnant. Written informed consent was obtained for each eligible participant. All participants were compensated $20.00 for their participation. This study was approved by the University of Alabama at Birmingham's Institutional Review Board.

Demographics, physical activity, and health conditions

Self-reported race/ethnicity, age, sex, medical history (health conditions and medication use), physical activity status, and T2D status were obtained through an interviewer-administered questionnaire. To assess physical activity status, participants were asked the question "What is your current activity level (i.e., person's average daily activity)?" and they responded by selecting one of five options: "none," "some," "moderate," "athlete," or "elite athlete." In this study, we grouped all participants into two physical activity groups: those who were not physically active (60.2%), which consisted of all individuals who reported either "none" or "some" regular physical activity, and the physically active group (39.8%), which consisted of participants who had "athlete" or "elite athlete" activity levels.

Body composition measurements

Body composition was assessed for all participants using two approaches. First, DXA (enCORE 2011 version 13.6; GE Lunar iDXA Corporation, Madison, Wisconsin) scans were used to estimate percentage body fat, denoted BF_{DXA}. Second, two-dimensional photographic images were processed for body volume and shape measures and were

then used to calculate body fat percentage, denoted BF_{Photo}. This programming algorithm has been described in more detail elsewhere (26). Photographic images were obtained using a digital camera (Canon PowerShot Model SX50; Canon USA Inc., Melville, New York). All participants wore close-fitting tank tops (females only) and spandex shorts for body composition measurements in order to reduce measurement bias. Trained staff measured weight to the nearest 0.1 kg using a physician's balance beam scale (Model 402LB; HealthOMeter, McCook, Illinois) and height to the nearest 0.1 cm using a stadiometer, and these measurements were used to calculate BMI for each participant.

Statistical methods

Descriptive statistics (mean ± SD) were calculated for the study sample. Mean body fat percentages (BF_{DXA} and BF_{Photo}) were assessed on the basis of physical activity and T2D status by one-way analysis of variance (ANOVA). Pearson correlation coefficients, *r*, were computed between pairs of model variables (e.g., age, height, weight, BF_{DXA}, BF_{Photo}). Linear regression analyses were performed to compare body fat measures. Using Bland-Altman analyses, we investigated the distribution of absolute and relative differences between body fat measures to assess any biases in BF_{Photo} relative to BF_{DXA} (29). Logistic regression analyses were performed to predict (1) the probability of having T2D and (2) the probability of being physically active. All statistical analyses were performed using SAS (version 9.4; SAS Institute, Cary, North Carolina) or R version 3.2 (R Development Core Team, Vienna, Austria), with statistical significance accepted when $P < 0.05$ (two-tailed).

Measurement error bias-adjustment methods

Three measurement error bias-adjustment methods were used: RC, SIMEX, and MI. Details of the methods and algorithms are summarized in Supporting Information Appendices S1-S3. The RC method can be applied to a validation study (internal and external), when a gold standard or imperfect reference instrument is available (8,19). However, for other studies, the noniterative RC method can be used to approximate regression coefficients from regression models with measurement error in covariates when a reference method is not available (see Carroll et al. and Spiegelman et al. for details) (8,19). The RC method consists of estimating model parameters of the logistic model with error-prone variables and covariates, estimating regression coefficients for a linear model that relates the error-free to the error-prone variables, and subsequently using the estimated parameters from the linear model to obtain the RC-adjusted model parameters. The 95% confidence interval (CI) for the regression coefficients and their respective odds ratios (OR) were calculated using the variance-covariance matrix for the bias-adjusted model parameters (10,19) (Supporting Information Appendix S1). RC was implemented using SAS macro %blinplus (7,10,19).

The SIMEX method, developed by Carroll et al. (14), is a simulation-based approach that reduces the bias in parameter estimates due to measurement error by introducing random error into the model (5,14,15,19). Simulated data with additive error terms were used to characterize the relation between model parameters and the amount of measurement error through a resampling approach. To characterize this trend, the parameter estimates were modeled as a function of the measurement error and corresponding mean regression coefficients. In the next step, the model parameter estimates for a model with error-free predictors were obtained by extrapolating back to the case of zero error (Supporting Information Appendix S2). This was implemented using the R "simex" package (15).

The MI framework was applied by treating the true values of the variables with measurement error as a missing data problem (5,11,12), i.e., imputing the bias-adjusted values for unadjusted values. Multiple (e.g., *m*) values or "imputations" were imputed for each unadjusted value under the MI principle and were used to replace the unadjusted value so that *m* data sets with only bias-adjusted values were generated. In the subsequent analysis, the *m* data sets were analyzed individually, yielding *m* statistics (e.g., mean, parameter estimates). Eventually, the *m* statistics were combined into a single statistic using Rubin's rule (30) (Supporting Information Appendix S3). The MI method has been implemented in many statistical software packages, such as SAS version 9.4, R version 3.2, and others (SPSS Statistics, version 24, IBM Corp., Armonk, New York (31); Stata, release 15, StataCorp LLC, College Station, Texas) so that bias adjustment by MI can also be utilized conveniently.

Results

Body composition assessment

The sample consisted of non-Hispanic white (51%) and non-Hispanic black (49%) adult men (46%) and women (54%) aged 39 ± 15 years (mean ± SD) with BMI of 28 ± 6 kg/m^2. About 5.3% of the participants self-reported having T2D, and 39.8% self-reported being physically active. Additional participant characteristics are summarized in Table 1. As expected, individuals considered to be physically active tended to have lower body fat percentages (Figure 1A). Body fat as estimated by BF_{DXA} and BF_{Photo} was significantly greater in the individuals who were not physically active (BF_{DXA}: 37% ± 10% and BF_{Photo}: 36% ± 10%) than in the physically active group (BF_{DXA}: 26% ± 9% and BF_{Photo}: 30% ± 7%; $P < 0.0001$). However, individuals with T2D had more variability in their body fat percentages (Figure 1B). Body fat estimates were significantly lower in

TABLE 1 Summary of participant characteristics

	All (*N* = 588)	Men (*n* = 270)	Women (*n* = 318)
Age (y)	39 ± 15[a]	40 ± 16	38 ± 14
Height (cm)	169.9 ± 9.8	177.3 ± 7.3	163.5 ± 6.7
Weight (kg)	80.0 ± 19.8	85.2 ± 19.2	75.7 ± 19.2
BMI (kg/m^2)	27.7 ± 6.3	27 ± 5.5	28.3 ± 6.9
BF_{DXA} (%)	32.9 ± 11.2	26.1 ± 8.9	38.8 ± 9.5
BF_{Photo} (%)	33.3 ± 9.7	26.6 ± 6.8	39.1 ± 7.9
Race, *n* (%)			
Non-Hispanic white	297 (50.5)	145 (53.7)	152 (47.8)
Non-Hispanic black	291 (49.5)	125 (46.3)	166 (52.2)
PA, *n* (%)			
No	354 (60.2)	136 (50.4)	218 (68.5)
Yes	234 (39.8)	134 (49.6)	100 (31.5)
T2D, *n* (%)			
No	557 (94.7)	255 (94.4)	302 (95.0)
Yes	31 (5.3)	15 (5.6)	16 (5.0)

Values are mean ± SD unless otherwise indicated.
PA, physically active.

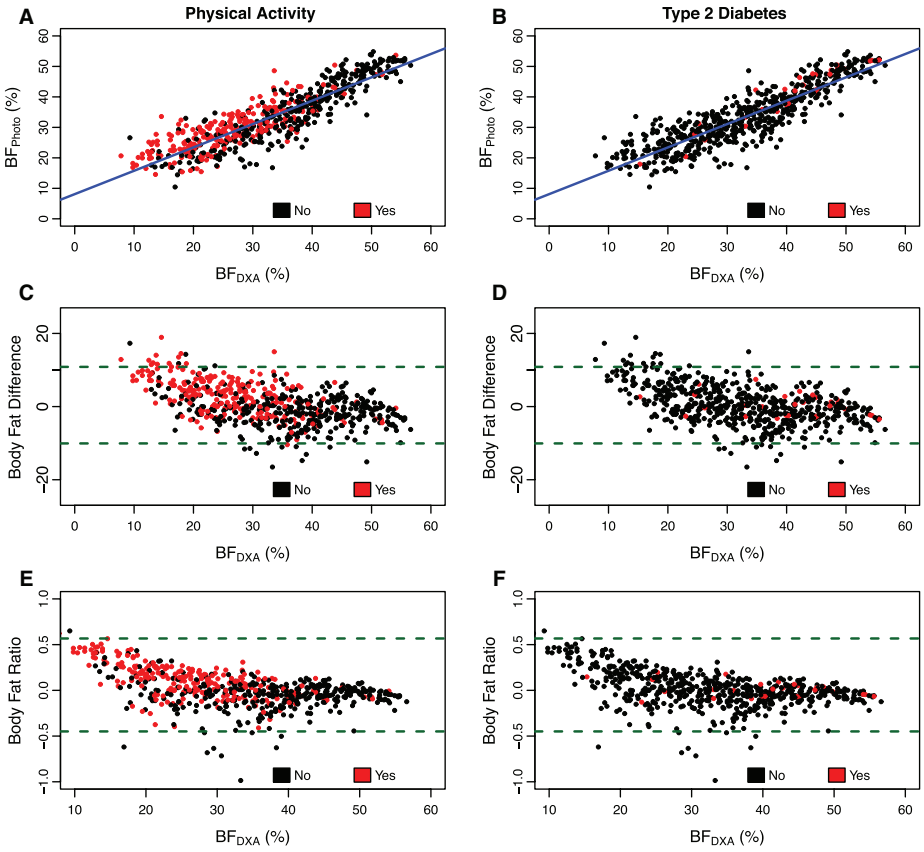

Figure 1 Associations and agreements between the two body fat percentage estimates. BF_{Photo} and BF_{DXA}, by (left column) physical activity and (right column) type 2 diabetes outcomes for all participants (n = 588). (A,B) BF_{Photo} strongly correlated with BF_{DXA} ($r = 0.88$, $P < 0.0001$). Bland-Altman tests with 95% CI (dashed green lines) for the (C,D) absolute difference $BF_{Photo} - BF_{DXA}$ and (E,F) relative difference $(BF_{Photo} - BF_{DXA}) / BF_{DXA}$ show that BF_{Photo} overestimated body fat percentage for individuals with lower body fat percentage.

TABLE 2 Correlation and variance-covariance summary

	Pearson correlation coefficients					Variance-covariance values				
	Age	**Height**	**Weight**	**BF_{DXA}**	**BF_{Photo}**	**Age**	**Height**	**Weight**	**BF_{DXA}**	**BF_{Photo}**
Age	1.00	−0.10	0.08*	0.22**	0.27**	228.61	−15.09	24.35	37.65	38.85
Height		1.00	0.41**	−0.42**	−0.44**		95.92	78.81	−46.24	−41.96
Weight			1.00	0.44**	0.49**			390.95	97.56	94.98
BF_{DXA}				1.00	0.88**				124.83	95.66
BF_{Photo}					1.00					93.85[1]

Correlation coefficients with *$P < 0.05$ and **$P < 0.0001$ are shown.

individuals with T2D (BF_{DXA}: 32% ± 11% and BF_{Photo}: 33% ± 10%) than in those who did not report T2D (BF_{DXA}: 40% ± 10% and BF_{Photo}: 40% ± 9%; $P < 0.0004$).

Pearson correlation coefficients (Table 2) indicated strong positive associations between BF_{DXA} and BF_{Photo} ($r = 0.88$, $P < 0.0001$; Figure 1). Bland-Altman analyses (Figure 1C-1F) were performed to evaluate the amount of bias in the BF_{Photo} data due to measurement error. The absolute mean difference ($\hat{u}_{abs}$) between BF_{Photo} and BF_{DXA} represents the average bias in body fat percentage. Here, it was $\hat{u}_{abs} = 0.38\%$ for all participants as a group (slope = −13.94; $P < 0.0001$), and the variance of this bias was $\hat{\sigma}^2_{abs} = 27.36$. The relative difference between BF_{Photo} and BF_{DXA} was 0.05% (95% CI: −0.44% to 0.56%).

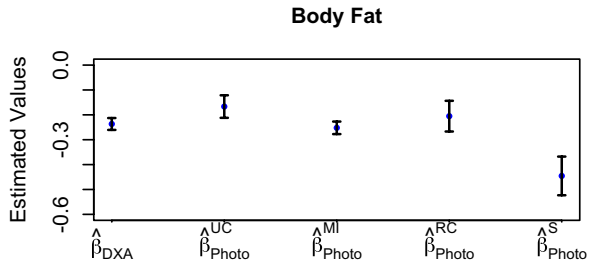

Figure 2 Physical activity status results. Parameters estimated with standard error bars for the probability of being physically active are shown for the error-free measurement (DXA), the unadjusted error-prone measurement (UC, Photo), and the three measurement error bias-adjusted cases (MI, RC, and S). DXA, dual-energy x-ray absorptiometry; UC, unadusted parameter; MI, multiple imputation; RC, regression calibration; S, simulation extrapolation. [Colour figure can be viewed at wileyonlinelibrary.com]

Measurement error bias adjustment

Parameters estimated from the logistic regression model with BF_{DXA} as the independent variable are denoted $\hat{\beta}_{DXA}$, and similarly, those estimated using BF_{Photo} are denoted $\hat{\beta}^{UC}_{Photo}$, where the superscript "UC" denotes an unadjusted parameter. Adjusted parameters of BF_{Photo} that are bias adjusted are denoted with appropriate superscripts, i.e., $\hat{\beta}^{MI}_{Photo}$ for the MI approach, $\hat{\beta}^{RC}_{Photo}$ for the RC method, and $\hat{\beta}^{S}_{Photo}$ for the SIMEX procedure.

Model coefficients for predicting the probability of being physically active are shown in Figure 2 and Table 3. The effect of BF_{DXA} on the odds of being physically active ($\hat{\beta}_{DXA} = -0.23$ [95% CI: −0.26 to −0.19]; OR 0.79 [95% CI: 0.75 to 0.82]; $P < 0.0001$) was greater than that of BF_{Photo} ($\hat{\beta}^{UC}_{Photo} = -0.16$ [95% CI: −0.23 to −0.08]; OR 0.84 [95% CI: 0.77 to 0.92]; $P = 0.0002$), which had an upward bias ($\hat{\beta}^{UC}_{Photo} > \hat{\beta}_{DXA}$). Similarly, an upward bias was observed for RC-adjusted estimates ($\hat{\beta}^{RC}_{Photo} > \hat{\beta}_{DXA}$), but a downward bias was found for SIMEX and MI ($\hat{\beta}^{S}_{Photo}, \hat{\beta}^{MI}_{Photo} < \hat{\beta}_{DXA}$). To quantify the improvement of the measurement error bias-adjustment methods, the percentage changes for the adjusted model parameters, denoted $\Delta\beta(\%)$, were calculated as the absolute change between the bias-adjusted parameter value (e.g., $\hat{\beta}^{UC}_{Photo}, \hat{\beta}^{MI}_{Photo}, \hat{\beta}^{RC}_{Photo}, \hat{\beta}^{S}_{Photo}$) and the reference parameter value ($\hat{\beta}_{DXA}$) divided by the absolute reference parameter value ($\hat{\beta}_{DXA}$). The percentage change was lowest for MI ($\Delta\beta = 9\%$), next lowest for RC ($\Delta\beta = 13\%$), and significantly higher for SIMEX ($\Delta\beta = 91\%$). Other significant covariates (adjusted and unadjusted) were age, sex, race, and weight (not shown); however, RC-adjusted age was not significant.

Model parameters estimated for predicting the probability of having T2D are shown in Table 4 and Figure 3. The effect of BF_{DXA} ($\hat{\beta}_{DXA} = 0.03$ [95% CI: −0.04 to 0.10]; OR 1.03 [95% CI: 0.95 to 1.12]; $P = 0.4447$) on the odds of having T2D was smaller in comparison to BF_{Photo} ($\hat{\beta}^{UC}_{Photo} = 0.10$ [95% CI: −0.05 to 0.25]; OR 1.11 [95% CI: 0.93 to 1.32]; $P = 0.2188$) and had an upward bias ($\hat{\beta}^{UC}_{Photo} > \hat{\beta}_{DXA}$). Both RC and SIMEX had an upward bias ($\hat{\beta}^{RC}_{Photo}, \hat{\beta}^{S}_{Photo} > \hat{\beta}_{DXA}$); however, MI was unbiased ($\Delta\beta = 0\%; \hat{\beta}^{MI}_{Photo} = \hat{\beta}_{DXA}$). The percentage change was more

TABLE 3 Summary of physical activity results (estimated model coefficients for body fat are shown for being physically active)

Model coefficient	Parameter[a]	Estimate (95% CI)	%$\Delta\beta$[b,c]	OR (95% CI)	*P*
Body fat	$\hat{\beta}_{DXA}$	−0.23 (−0.26 to −0.19)	–	0.79 (0.75 to 0.82)	<0.0001
	$\hat{\beta}^{UC}_{Photo}$	−0.16 (−0.23 to −0.08)	30 ↑	0.84 (0.77 to 0.92)	0.0002
	$\hat{\beta}^{MI}_{Photo}$	−0.25 (−0.28 to −0.21)	9 ↓	0.77 (0.73 to 0.81)	<0.0001
	$\hat{\beta}^{RC}_{Photo}$	−0.20 (−0.31 to −0.08)	13 ↑	0.81 (0.72 to 0.91)	0.0008
	$\hat{\beta}^{S}_{Photo}$	−0.44 (−0.57 to −0.30)	91 ↓	0.64 (0.55 to 0.74)	<0.0001

[a]Unadjusted BF_{Photo} ($\hat{\beta}^{UC}_{Photo}$), multiple imputation ($\hat{\beta}^{MI}_{Photo}$), regression calibration ($\hat{\beta}^{RC}_{Photo}$), and SIMEX ($\hat{\beta}^{S}_{Photo}$).
[b]Difference between $\hat{\beta}_{DXA}$ and each parameter value (e.g., $\hat{\beta}^{UC}_{Photo}, \hat{\beta}^{MI}_{Photo}, \hat{\beta}^{RC}_{Photo}, \hat{\beta}^{S}_{Photo}$). Percentage change, denoted **$\Delta\beta$(%)**, calculated as absolute change between BF_{Photo}-based parameter value (e.g., $\hat{\beta}^{UC}_{Photo}, \hat{\beta}^{MI}_{Photo}, \hat{\beta}^{RC}_{Photo}, \hat{\beta}^{S}_{Photo}$) and reference parameter value ($\hat{\beta}_{DXA}$), divided by absolute reference parameter value ($\hat{\beta}_{DXA}$), and rounded.
[c]↑ represents upward bias ($\hat{\beta} > \hat{\beta}_{DXA}$), and ↓ represents downward bias ($\hat{\beta} > \hat{\beta}_{DXA}$).

TABLE 4 Summary of T2D results (estimated model coefficients of body fat are shown for assessing T2D status)

Model coefficient	Parameter[a]	Estimate (95% CI)	%$\Delta\beta$[b,c]	OR (95% CI)	*P*
Body fat	$\hat{\beta}_{DXA}$	0.03 (−0.04 to 0.10)	–	1.03 (0.95 to 1.12)	0.4447
	$\hat{\beta}^{UC}_{Photo}$	0.10 (−0.05 to 0.25)	233 ↑	1.11 (0.93 to 1.32)	0.2188
	$\hat{\beta}^{MI}_{Photo}$	0.03 (−0.02 to 0.08)	0	1.03 (0.96 to 1.10)	0.3724
	$\hat{\beta}^{RC}_{Photo}$	0.13 (−0.06 to 0.32)	333 ↑	1.14 (0.92 to 1.41)	0.2245
	$\hat{\beta}^{S}_{Photo}$	0.27 (−0.02 to 0.56)	800 ↑	1.32 (0.98 to 1.77)	0.0653

[a]Unadjusted BF_{Photo} ($\hat{\beta}^{UC}_{Photo}$), multiple imputation ($\hat{\beta}^{MI}_{Photo}$), regression calibration ($\hat{\beta}^{RC}_{Photo}$), and SIMEX ($\hat{\beta}^{S}_{Photo}$).
[b]Difference between $\hat{\beta}_{DXA}$ and each parameter value (e.g., $\hat{\beta}^{UC}_{Photo}, \hat{\beta}^{MI}_{Photo}, \hat{\beta}^{RC}_{Photo}, \hat{\beta}^{S}_{Photo}$). Percentage change, denoted **$\Delta\beta$(%)**, calculated as absolute change between BF_{Photo}-based parameter value (e.g., $\hat{\beta}^{UC}_{Photo}, \hat{\beta}^{MI}_{Photo}, \hat{\beta}^{RC}_{Photo}, \hat{\beta}^{S}_{Photo}$) and reference parameter value ($\hat{\beta}_{DXA}$), divided by absolute reference parameter value ($\hat{\beta}_{DXA}$), and rounded.
[c]↑ represents upward bias ($\hat{\beta} > \hat{\beta}_{DXA}$), and ↓ represents downward bias ($\hat{\beta} > \hat{\beta}_{DXA}$).

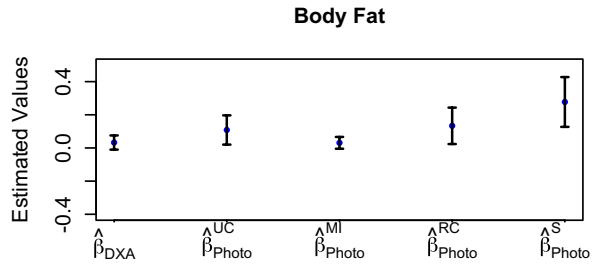

Figure 3 Type 2 diabetes status results. Parameters estimated with standard error bars for the probability of having type 2 diabetes are shown for the error-free measurement (DXA), the unadjusted error bias-adjusted measurement (UC, Photo), and the three measurement error bias-adjusted cases (MI, RC, and S). DXA, dual-energy x-ray absorptiometry; MI, multiple imputation; RC, regression calibration; S, simulation extrapolation. [Colour figure can be viewed at wileyonlinelibrary.com]

than 300% for RC ($\Delta\beta = 333\%$) and SIMEX ($\Delta\beta = 800\%$). Parameter estimates for the age variable (adjusted and unadjusted) were the only significant predictor for T2D status (not shown), with the exception of SIMEX-adjusted age.

Discussion

We presented and compared three measurement error bias-adjustment methods, RC, MI, and SIMEX, to reduce potential biases in statistical models used to evaluate the effect of body fat on health outcomes. The performances of these commonly used measurement error techniques were compared using body fat percentage estimated by DXA scans (BF_{DXA}) as the reference measure and a novel two-dimensional photographic-based method (BF_{Photo}) as the error-prone measure. We applied the error bias-adjustment methods to a logistic model involving body fat to predict the probability of two health outcomes: having T2D and being physically active.

In this biethnic sample of adults, BF_{Photo} and BF_{DXA} measures were strongly positively correlated; however, BF_{Photo} exhibited bias as measured by Bland-Altman analyses (Figure 1). Body fat percentage (BF_{Photo} and BF_{DXA}) was lower in individuals with higher physical activity levels compared with individuals who were not active and was a significant predictor of being physically active (Figure 1), which was expected on the basis of prior studies of body composition and physical activity levels (32,33). This association of body fat with the probability that an individual is physically active was also observed when the population was stratified by race. In contrast, BF_{Photo} and BF_{DXA} were lower in participants with self-reported T2D relative to individuals without T2D, which was not expected. That is, contrary to our expectations, excess body fat did not significantly increase the odds of having T2D in our cohort (Table 4). Furthermore, the association of body fat with the probability of having T2D when the population was stratified by race was inconsistent, in that it was statistically significant in some cases and not significant in others. However, the parameter estimates and 95% CI led to the conclusion that the association of body fat remained null, which is consistent with the findings shown in this study when the sample was analyzed as one group. Therefore, it is possible that our study participants with T2D had better body weight management than average: all participants with T2D reported being on medication(s) used for diabetes, weight loss, or other health conditions, which may explain the inverse relationship between body fat percentage and having T2D. The lower body fat percentage observed in participants with T2D could also be due to the cross-sectional design of the study, as well as the particular sample that volunteered to enroll in the study.

We also observed an upward bias from unadjusted BF_{Photo} data ($\hat{\beta}^{UC}_{Photo} > \hat{\beta}_{DXA}$) for models predicting the probability of being physically active and having T2D, which implies that unadjusted BF_{Photo} overestimated the effect of body fat on health outcomes. The performance of measurement error bias-adjustment methods to reduce this bias varied. For the physical activity model, SIMEX and MI methods led to a downward bias ($\hat{\beta}^{S}_{Photo}, \hat{\beta}^{MI}_{Photo} < \hat{\beta}_{DXA}$), whereas the RC-adjusted estimate had an upward bias ($\hat{\beta}^{RC}_{Photo} > \hat{\beta}_{DXA}$); however, MI had the lowest percent change ($\Delta\beta = 9\%$). Thus, MI performed the best, followed next by RC, and then the SIMEX method. For the T2D status model, both RC-adjusted and SIMEX-adjusted values had an upward bias ($\hat{\beta}^{RC}_{Photo}, \hat{\beta}^{S}_{Phpto} > \hat{\beta}_{DXA}$), whereas the MI-adjusted estimate was unbiased ($\hat{\beta}^{MI}_{Photo} = \hat{\beta}_{DXA}$) and therefore performed the best. The RC- and SIMEX-adjusted model parameters corresponding to body fat had percentage changes exceeding 300% and overestimated the effect of body fat percentage on the probability of having T2D. Therefore, these results indicate that parameter estimates bias-adjusted by MI were closer to the reference estimates in comparison to RC and SIMEX. More specifically, MI-adjusted estimates consistently matched the estimates corresponding to BF_{DXA} for both T2D and physical activity outcomes. RC overestimated the effect of body fat ($\hat{\beta}^{RC}_{Photo}$) on the probability of having T2D but improved the estimate of the effect of body fat ($\hat{\beta}^{RC}_{Photo}$) on the probability of being physically active. However, SIMEX overestimated the effect of body fat ($\hat{\beta}^{S}_{Photo}$) for both T2D and physical activity status outcomes. The standard error was greater for RC and SIMEX than for MI. Moreover, similar findings on the performance of MI, RC, and SIMEX for the bias adjustment of parameter estimates have been reported in other studies (5,11,12,34).

An advantage of our research is that we have concurrent body fat measures by DXA and a two-dimensional photographic-based method for the entire sample. Another novel aspect of our work is the use of measurement error approaches to improve model parameters of body composition, which, to our knowledge, has not been done before. A limitation of this work is that the three error bias-adjustment methods that we implemented have different assumptions, which makes comparison of their performance difficult. For example, whereas MI may be the most effective method for measurement error bias adjustment in this study, the imputation procedure includes T2D and physical activity outcome variables to simulate new data sets, thus improving its accuracy. In contrast to MI, the SIMEX and RC methods are completely different approaches with a different set of assumptions (see online Supporting Information Appendices S1-S3). Furthermore, it is important to note that the RC method can be applied to validation studies (internal and external), when a gold standard or reference measure is available (8,19). However, in cases when this is not available, the noniterative RC method can be used to adjust for measurement error in covariates (see Carroll et al. and Spiegelman et al. for details) (8,19). Another limitation is that these methods cannot address other biases, such as those caused by unmeasured confounders and other biases. The results in our analyses are sensitive to all types of errors, not just measurement error but also unmeasured confounders (35,36). While these methods cannot address bias caused by unmeasured confounders, the measurement error models discussed here provide a valuable

method for addressing measurement error when data are available and measurement error is a concern. Future work would consider unmeasured confounders when adjusting for measurement error. Lastly, another limitation is that the self-reported outcomes considered in this study, T2D and physical activity, may potentially be misclassified and therefore could affect our results. However, the focus of this study is to assess the impacts of measurement error in the covariates on these outcomes. Future work would involve exploring the effects of errors on the outcomes.

In conclusion, this study demonstrates the value of measurement error bias-adjustment methods to improve model parameters in nutrition and obesity research studies. Our purpose was to introduce three statistical approaches for reducing bias due to measurement errors and to illustrate their value using real data. We presented a practical example that involves evaluating the relationship between body fat and health outcomes. These tools were applied for a specific statistical model (logistic model) and data set (body fat measured by DXA and a novel photographic-based method). Our results suggest that, overall, MI performed the best in adjusting for measurement error and can be used to minimize statistical bias caused by measurement error, a finding that is supported by other studies (11,12). Furthermore, our study evaluated the case in which the variance in the reference measure (DXA) was larger than the error-prone method (photographic), and here, MI clearly outperformed all other methods. In summary, this study illustrates the utility and value of these methods for investigators conducting research where measurement error is a concern.O

Acknowledgments

The authors would also like to acknowledge two anonymous reviewers for their careful review and suggestions on the research aims and study findings, which have improved the quality of, as well as the clarity of the topics discussed in, the manuscript.

References

1. Horner NK, Patterson RE, Neuhouser ML, Lampe JW, Beresford SA, Prentice RL. Participant characteristics associated with errors in self-reported energy intake from the Women's Health Initiative food-frequency questionnaire. *Am J Clin Nutr* 2002;76:766-773.
2. Prentice RL. Measurement error and results from analytic epidemiology: dietary fat and breast cancer. *J Natl Cancer Inst* 1996;88:1738-1747.
3. Murray RP, Connett JE, Lauger GG, Voelker HT. Error in smoking measures: effects of intervention on relations of cotinine and carbon monoxide to self-reported smoking. The Lung Health Study Research Group. *Am J Public Health* 1993;83:1251-1257.
4. Duffy JC, Alanko T. Self-reported consumption measures in sample surveys: a simulation study of alcohol consumption. *J Off Stat* 1992;8:327-350.
5. Freedman LS, Midthune D, Carroll RJ, Kipnis V. A comparison of regression calibration, moment reconstruction and imputation for adjusting for covariate measurement error in regression. *Stat Med* 2008;27:5195-5216.
6. Prentice RL, Huang Y, Kuller LH, et al. Biomarker-calibrated energy and protein consumption and cardiovascular disease risk among postmenopausal women. *Epidemiology* 2011;22:170-179.
7. Rosner B, Spiegelman D, Willett W. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. *Am J Epidemiol* 1990;132:734-745.
8. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. *Measurement Error in Nonlinear Models: A Modern Perspective*. Boca Raton, FL: CRC Press; 2006.
9. Fuller WA. *Measurement Error Models*. Hoboken, NJ: John Wiley & Sons; 2009.
10. Spiegelman D, McDermott A, Rosner B. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. *Am J Clin Nutr* 1997;65:1179S-1186S.
11. Cole SR, Chu H, Greenland S. Multiple-imputation for measurement-error correction. *Int J Epidemiol* 2006;35:1074-1081.
12. Padilla MA, Divers J, Vaughan LK, Allison DB, Tiwari HK. Multiple imputation to correct for measurement error in admixture estimates in genetic structured association testing. *Hum Hered* 2009;68:65-72.
13. Messer K, Natarajan L. Maximum likelihood, multiple imputation and regression calibration for measurement error adjustment. *Stat Med* 2008;27:6332-6350.
14. Carroll RJ, Küchenhoff H, Lombard F, Stefanski LA. Asymptotics for the SIMEX estimator in nonlinear measurement error models. *J Am Stat Assoc* 1996;91:242-250.
15. Lederer W, Küchenhoff H. A short introduction to the SIMEX and MCSIMEX. *R News*. October 2006:26-31. https://cran.r-project.org/doc/Rnews/Rnews_2006-4.pdf.
16. Guolo A. Robust techniques for measurement error correction: a review. *Stat Methods Med Res* 2008;17:555-580.
17. Thürigen D, Spiegelman D, Blettner M, Heuer C, Brenner H. Measurement error correction using validation data: a review of methods and their applicability in case-control studies. *Stat Methods Med Res* 2000;9:447-474.
18. Thomas D, Stram D, Dwyer J. Exposure measurement error: influence on exposure-disease relationships and methods of correction. *Annu Rev Public Health* 1993;14:69-93.
19. Spiegelman D, Carroll RJ, Kipnis V. Efficient regression calibration for logistic regression in main study/internal validation study designs with an imperfect reference instrument. *Stat Med* 2001;20:139-160.
20. Prentice RL, Shaw PA, Bingham SA, et al. Biomarker-calibrated energy and protein consumption and increased cancer risk among postmenopausal women. *Am J Epidemiol* 2009;169:977-989.
21. Rothman KJ. BMI-related errors in the measurement of obesity. *Int J Obes* 2008;32:S56-S59.
22. Ulijaszek SJ, Kerr DA. Anthropometric measurement error and the assessment of nutritional status. *Br J Nutr* 1999;82:165-177.
23. Chiu C-Y, Sanders RH. Quantifying obesity from anthropometric measures and body volume data. *International Journal of Design, Analysis and Tools for Integrated Circuits and Systems* 2009;1:1-4. http://shura.shu.ac.uk/17742/
24. Lee SY, Gallagher D. Assessment methods in human body composition. *Curr Opin Clin Nutr Metab Care* 2008;11:566-572.
25. Garlie TN, Obusek JP, Corner BD, Zambraski EJ. Comparison of body fat estimates using 3D digital laser scans, direct manual anthropometry, and DXA in men. *Am J Hum Biol* 2010;22:695-701.
26. Pradhan L, Song G, Zhang C, et al. Feature extraction from 2D images for body composition analysis. Published in: *IEEE ISM 2015*; Dec 14-16, 2015; Miami, FL:42-52.
27. Affuso O, Pradhan L, Zhang C, et al. A method for measuring human body composition using digital images. *PLoS One* 2018:13:e0206430. doi:10.1371/journal.pone.0206430
28. Capers PL, Kinsey AW, Miskell EL, Affuso O. Visual representation of body shape in African-American and European American women: clinical considerations. *Clin Med Insights Womens Health* 2016;9:63-70.
29. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135-160.
30. Li P, Stuart EA, Allison DB. Multiple imputation: a flexible tool for handling missing data. *JAMA* 2015;314:1966-1967.
31. IBM Corporation. IBM SPSS Missing Values 24. Armonk, NY: IBM Corp; 2016. ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/24.0/en/client/Manuals/IBM_SPSS_Missing_Values.pdf. Accessed January 12, 2018.
32. Zanovec M, Lakkakula AP, Johnson LG, Turri G. Physical activity is associated with percent body fat and body composition but not body mass index in white and black college students. *Int J Exercise Sci* 2009;2:175-185.
33. Bowen L, Taylor AE, Sullivan R, et al. Associations between diet, physical activity and body fat distribution: a cross sectional study in an Indian population. *BMC Public Health* 2015;15:281. doi:10.1186/s12889-015-1550-7
34. Divers J, Vaughan LK, Padilla MA, Fernandez JR, Allison DB, Redden DT. Correcting for measurement error in individual ancestry estimates in structured association tests. *Genetics* 2007;176:1823-1833.
35. Goldsmith K, Chalder T, White P, Sharpe M, Pickles A. Measurement error, time lag, unmeasured confounding: considerations for longitudinal estimation of the effect of a mediator in randomised clinical trials. *Stat Methods Med Res* 2018;27:1615-1633.
36. Fewell Z, Davey Smith G, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *Am J Epidemiol* 2007;166:646-655.