Obesity Science & Practice doi: 10.1002/osp4.300

ORIGINAL ARTICLE

External Validation of Equations that Use Demographic and Anthropometric Measurements to Predict Percent Body Fat

K. R. Reynolds[1], J. Stevens[1,2], J. Cai[3], C. E. Lewis[4], A. C. Choh[5] and S. A. Czerwinski[5]

[1]Departments of Epidemiology, University of North Carolina, Chapel Hill, NC, US; [2]Departments of Nutrition, University of North Carolina, Chapel Hill, NC, US; [3]Departments of Biostatistics, University of North Carolina, Chapel Hill, NC, US; [4]Division of Preventive Medicine, University of Alabama at Birmingham, Birmingham, AL, US; [5]School of Public Health, University of Texas Health Science Center at Houston, Brownsville, TX, US.

Received 25 May 2018; revised 23 August 2018; accepted 26 August 2018

Correspondence:
June Stevens, Departments of Epidemiology, University of North Carolina, Chapel Hill, NC, US, CB 7461, University of North Carolina, Chapel Hill, North Carolina, 27599.
E-mail: june_stevens@unc.edu

Summary

Objective

Numerous equation to predict percent body fat using demographics and anthropometrics have been published but external validation of these equations is limited. The objective of this study was to validate published equations that use anthropometrics for prediction of percent body fat using external data.

Methods

Data were from the Visceral Fat, Metabolic Rate, and Coronary Heart Disease Risk I (VIM I) Study and the Fels Longitudinal Study (Fels). VIM I was conducted in a subset of subjects from the CARDIA study and included black and white adults 28–40 years (n = 392). Fels consisted of white participants 8–88 years (n = 1,044). Percent body fat assessed by dual X-ray absorptiometry (DXA) in these two studies was compared to results calculated using 13 equations from Stevens et al. and nine other published equations.

Results

In general, the Stevens equations performed better than equations from other studies. For example, equation “I“ in women in VIM I, Fels adults, and Fels youth, R^2 estimates were 0.765, 0.757 and 0.789, respectively. In men the estimates were 0.702 in VIM I, 0.822 in Fels adults and 0.905 in Fels youth. None of the results from the nine published equations showed R^2 this high in corresponding groups.

Conclusions

Our results indicate that several of the Stevens equations have external validity superior to that of nine other published equations among varying age groups, genders and races.

Keywords: Adiposity, body composition, DXA.

Introduction

Equations that include demographics and anthropometrics support a feasible method for assessing body composition in a wide range of settings. To develop such equations, it is a common practice to divide the data into two parts: one used to select variables and calculate coefficients against a criterion measure of fatness such as DXA, and the other used to provide internal validation of the equations developed. If the performance of the equation in the internal validation sample is substantially worse than in the development sample, it may indicate overfitting and poor generalizability. It is likely that equations perform more strongly in internal validation than when applied to a different, or external study because study staff, measurement protocols and study participants are almost certain to be more similar within a single study. Since the main value of prediction equations for percent body fat (%BF) lies in their application to external study samples, it is important to establish the validity of the equations in samples other than those used for development.

Recently Stevens et al. created equations that used demographic and anthropometric variables for the prediction of %BF in a diverse group of participants aged 8 years and older from the 1999–2006 National Health and Nutrition Examination Survey (NHANES) (1). Fourteen sex-specific equations were developed in this nationally representative data, and shown to have low bias across race/ethnicity (white, black, Mexican-American, other), age (youth versus adults) and BMI categories (1). Stevens and colleagues tested the equations in a random sample of data from the same study that was not used in equation development and showed high internal validity. Nevertheless, the equations have not been tested in external data. The purpose of this study is to evaluate the external validity of 13 of the 14 sex-specific Stevens et al. equations for the prediction of %BF. Several other published equations are also tested for external validity to provide comparison and perspective. In all, the external validity of 22 published equations will be examined when applied to two extant data sets by comparing predicted %BF values to assessments by DXA.

Methods

DXA and anthropometry data from two studies

Two data sources were identified that included both DXA-measured %BF (the criterion method) and multiple anthropometric variables: the Visceral Fat, Metabolic Rate, and Coronary Heart Disease Risk I Study (VIM I) and the Fels Longitudinal Study (Fels) (2). VIM I is an ancillary study of the Coronary Artery Risk Development in Young Adults (CARDIA) study, which is an ongoing prospective cohort study initiated in 1985–86 (3). CARDIA recruited black and white young adult men and women, aged 18–30 years, from four cities in the United States with the intent of studying the development of cardiovascular risk (3). In 1995–96, VIM I drew participants from two CARDIA sites (Birmingham, AL; and Oakland, CA), and 397 participants were selected such that age, race (black or white), sex, and BMI were approximately balanced to represent the previous CARDIA examination (1992–93) (4). Five VIM I participants were ineligible for our study because they were pregnant, breastfeeding, gave birth in the 6 months prior to the DXA scan, weighed 300 lbs. or more, had a height greater than 6 ft. 5 in., or had hands or feet that were not scanned. After excluding participants with missing data ($n = 5$), the analytic sample consisted of 195 men and 197 women.

Fels is an ongoing prospective cohort of predominantly white Americans founded in 1929 in Yellow Springs, Ohio (2). At its inception, the mission of the Fels Longitudinal Study was to examine human growth and development in youth, and it has since expanded to examine the aetiology of chronic diseases. Fels is a restricted cohort in which family members of the original cohort are enrolled either *in utero* or after marriage. There were 1,122 Fels participants 8 to 88 years old who had visits in the years between 1999 and 2006. Among these, participants were excluded if they exceeded DXA table limits, had measurements taken at less than 8 years of age, or were missing needed variables. If a participant had repeated measures, one was selected for inclusion at random. Nine hundred and ninety-two participants were included in the Fels youth and adult datasets. The final analytic samples consisted of 350 men and 426 women aged 20 to 88 years in the Fels adult dataset; and 149 boys and 145 girls aged 8 to 19 years in the Fels youth dataset.

Total %BF was measured in VIM I and in Fels using the Hologic 2000 densitometer and the Hologic QDR 4500A (Hologic Inc., Bedford MA), respectively. As per the Schoeller et al. protocol, the correction factor of 5% was used with the Hologic QDR4500A measurements (5). Triceps and subscapular skinfolds thickness, waist and arm circumferences, standing height, and weight were measured in VIM I. Triceps and subscapular skinfolds thicknesses; waist, calf, thigh, and arm circumferences; standing height; and weight were measured in Fels. Details of the methods used to collect anthropometric methods in the VIM I and Fels studies are available on their respective websites (6–8).

Twenty-two published equations from six studies

Table 1 summarizes the equations from the literature that met our criteria for study: (1) variables used were found in the NHANES data, (2) the age distribution of the participants used overlapped the age range of Fels or VIM I, (3) the race of the participants used was the same as within Fels or VIM I (white or black), (4) data were not from Fels or VIM I, and (5) anthropometric variables were not restricted to height and weight. Stevens and colleagues published 14 sex-specific equations, 13 of which could be examined here. Stevens' equation A could not be validated because it included arm and leg lengths that were not available in either VIM I or Fels. Gender-specific equations published by Durnin and Womersley, Lean et al., Zanovec et al., Hassager et al., Kagawa et al., and Slaughter et al (9–14). If the authors created different equations for participants with different characteristics, we applied the appropriate equation to each subgroup. In cases in which multiple equations in one paper met our eligibility criteria and were developed in the same subjects, we selected for study the equation shown by the authors to have the best performance. Equations were examined in age and race subgroups only if there was representation of that subgroup in the sample used for equation development. Only the Stevens and colleagues and the Zanovec et al. publications identified samples of adults who were not white that were large

Table 1 Characteristics of published formulas used to evaluate percent body fat in the present study

Author/Year	N	Race/Ethnicity	Age (years)	Reference Methods	Equation Label	Variables included in Equations	Stratification for Sex
Stevens 2015	11 884 males or 9215 females USA	White Black Mexican – American Other	≥ 8	DXA- Hologic QDR 4500A	B	Base + triceps skinfold, subscapular skinfold, waist circumference, calf circumference, arm circumference, thigh circumference	S
					C	Base + subscapular skinfold, waist circumference, calf circumference, arm circumference, thigh circumference	S
					D	Base + triceps, waist circumference, calf circumference, arm circumference, thigh circumference	S
					E	Base + waist circumference, calf circumference, arm circumference, thigh circumference	S
					F	Base + triceps skinfold, subscapular skinfold, waist circumference, calf circumference	S
					G	Base + triceps skinfold, subscapular skinfold, waist circumference, arm circumference	S
					H	Base + triceps skinfold, subscapular skinfold, arm circumference	S
					I	Base + triceps skinfold, waist circumference	S
					J	Base + waist circumference, arm circumference	S
					K	Base + arm circumference	S
					L	Base + waist circumference	S
					M	Base + triceps skinfold	S
					N	Base	S
Durnin and Womersley 1973	209 males 272 females United Kingdom	Unknown	16–72	UWW	1	triceps skinfold, subscapular skinfold (age spec.)	S
					2	triceps skinfold, subscapular skinfold (general)	S
Kagawa 2008	95 males 121 females Australia	Unknown	≥20	DXA -Lunar DPX-L	3	waist to height ratio at waist circumference, age, sex	N
Zanovec 2012	5,981 males 5,926 females USA	White Black Mexican - American	≥20	DXA - Hologic QDR 4500A	4	waist circumference, age, race/ethnicity, sex	N
Lean 1996	63 males 84 females United Kingdom	White	16–65	UWW	5	body mass index, triceps skinfold, age	S
					6	mid upper arm circumference, triceps skinfold, age	S
					7	waist circumference, triceps skinfold, age	S
Hassager 1986	98 males 130 females Denmark	Unknown	20–72	DPA	8	triceps skinfold, subscapular skinfold	S
Slaughter 1988[a]	310 participants USA	White Black	8–18	UWW	9	triceps skinfold, subscapular skinfold	S

S, sex stratified; N, not sex stratified; %BF, Percent Body Fat; BD, Body Density

[a] Equations were developed in a sample of 8 to 29 year olds but are recommended for predicting body fat in 8 to 18 year olds

Base: age, race, height, weight, body mass index.

enough for separate study, and only Stevens et al. and Slaughter et al. examined youth.

Descriptive information pertinent to the equations is shown in Table 2 with the Stevens equations labelled with letters (as was done in Stevens et al., 2015) and each published equation arbitrarily assigned a number from 1 to 9 (1). These letter and numbers are used here to facilitate equation identification across tables and figures. Samples used for equation development varied in size from 147 to 11,907 and were from the United States, Europe, or Australia. Criterion methods for published equation development included underwater weighing (UWW), dual photon absorptiometry (DPA), and dual x-ray absorptiometry (DXA).

Statistical analyses

All analyses were performed using SAS 9.4 (SAS Institute, Cary, NC). Percent body fat was calculated using the equations described by each author. A SAS macro available at used to calculate %BF predicted values for the Stevens equations (http://sph.unc.edu/nutr/american-body-comp-calculator/). The PROC REG procedure was used to model sex, race and/or age specific univariate linear regression models with DXA-measured %BF as the outcome and calculated %BF as the exposure. Root mean square errors (RMSE) and mean signed differences (MSD) were calculated to examine non-differential and differential error. To guide evaluation, we arbitrarily called it a weakness in performance if the R^2 was <0.75, the RMSE >4.5 or the MSD outside the bounds of -2 and 2.

Results

Demographic and anthropometric characteristics of the Fels and VIM I samples are shown in Table 2 overall and stratified by race or age. The mean age of the adults in Fels was about 12 years older than those in VIM I. The distributions of height (cm), weight (lbs), triceps skinfold thickness (mm), and thigh circumference (cm) were similar in the two adult samples. BMI was categorized into normal weight (18.5–24.9 kg/m^2), overweight (25.0–29.9 kg/m^2), and obese ($\geq$30 kg/m^2) in adults and overweight and obesity were defined as $>$85th BMI percentile and $>$ 95th BMI percentile in youth. The mean BMI was in the overweight range with more than 50% of the population having overweight or obesity. In VIM I, the mean

Table 2 Characteristics of the VIM I and Fels study participants

	VIM I						Fels					
	Overall		Black (Age 28-40)		White (Age 28-40)		Overall		Adults (Age 20 - 88)		Youth (Age 8 - 19)	
	N=392		*N*=191		*N*=201		*N*=992		*N*=776		*N*=294	
	%	Mean ± SD	%	Mean ± SD	%	Mean ± SD	%	Mean ± SD	%	Mean ± SD	%	Mean ± SD
Demography												
Age		35.1 ± 3.6		34.4 ± 3.7		35.8 ± 3.4		39.0 ± 21.0		46.8 ± 17.0		13.8 ± 3.5
Sex												
Male	50		52		48		47		45		51	
Female	50		48		52		53		55		49	
Race												
Black	49		100		0		0		0		0	
White	51		0		100		100		100		100	
BMI Category												
Normal	41		29		53		50		44		69	
Overweight	36		38		33		30		36		12	
Obese	22		32		12		19		19		15	
Anthropometry												
Arm Circumference (cm)		31.6 ± 4.4		32.8 ± 4.3		30.4 ± 4.2		30.5 ± 5.2		32.1 ± 4.1		26.1 ± 5.0
Height (cm)		171.5 ± 9.4		171.1 ± 9.5		172.0 ± 9.4		167.4 ± 13.4		171.3 ± 9.5		158.9 ± 17.0
Weight(kg)		78.6 ± 15.5		82.5 ± 15.1		74.9 ± 15.1		70.6 ± 20.1		76.8 ± 16.1		54.6 ± 19.3
BMI(kg/m^2)		26.6 ± 4.9		28.1 ± 5.0		25.2 ± 4.3		24.8 ± 5.2		26.1 ± 4.5		20.9 ± 4.6
Triceps skinfold(mm)		18.7 ± 9.8		19.8 ± 10.8		17.7 ± 8.6		17.8 ± 7.9		18.7 ± 7.9		15.1 ± 7.2
Subscapular Skinfold(mm)		19.1 ± 9.7		22.3 ± 10.5		16.1 ±7.6		17.8 ± 8.1		19.7 ± 7.3		12.2 ± 7.1
Waist Circumference (cm)		84.4 ± 11.4		86.8 ± 10.9		82.1 ± 11.5		89.8 ± 15.9		94.4 ± 13.1		76.2 ± 13.7
Calf Circumference(cm)		N/A		N/A		N/A		36.1 ± 4.3		37.2 ± 3.4		33.15 ± 4.94
Thigh Circumference(cm)		51.9 ± 5.7		53.3 ± 5.9		50.6 ± 5.3		49.4 ± 6.7		51.1 ± 5.2		45.3 ± 7.7
DXA %BF		30.4 ± 10.8		30.6 ± 12.0		30.3 ± 9.5		31.3 ± 8.3		32.2 ± 8.1		27.9 ± 8.1

DXA-measured %BF was 30.4% and in Fels adults, it was 32.2%. Compared to white participants, black participants in VIM I tended to have higher levels of all anthropometrics except height. The mean age of youth in Fels was 14 years, and approximately 27% of participants had overweight or obesity, and the mean %BF was 27.9%.

Seven of the Stevens et al. equations and nine of the other equations were examined for model fit in VIM I men and women; however, only the equations from Stevens et al. and from Zanovec et al. (equation #4) could be applied to black participants. Therefore, in Tables 3 (women) and 4 (men) there are fewer estimates shown for black participants alone and for black and white participants combined than for white participants. Table 3 shows that in black and white women combined and in white women alone, R^2 was over 0.75 in three of 7 Stevens et al. equations tested in VIM I and in 8 of 13 equations tested in Fels. None of the equations from the five other studies had an R^2 this high. The RSME was lower for the equations with a higher R^2. Among the Stevens et al. equations, equation I performed well in both the VIM I and the Fels data with fewer variables compared to other equations with similar performance. RMSE estimates tended to be lower in Fels women than VIM I women and tended to be lower in black women compared to white women in VIM I. In black women the Stevens et al. equations all had R^2 less than 0.700 with most below 0.600 and the Zanovec et al. (equation #4) produced a smaller R^2 (0.400) than any of the Stevens et al. equations. RMSE was consistently lower in women from Fels than from VIM I. In Fels R^2 estimates for the Stevens et al. equations tended to be higher in girls than women, but there were exceptions (equations E, K and N), and there were no strong trends in the comparisons of RMSE in the Fels youth versus adults. The Slaughter et al. equation (equation #9) had an R^2 of 0.745 and an RMSE of 3.40.

In males (Table 4), only equation I had an R^2 that was over 0.75 in VIM I, but several were over 0.9 in Fels youth. Among the other equations, only the Slaughter et al. equation (equation #9) applied to youth had an R^2 higher than 0.75. In contrast to what was observed in women from VIM I, R^2 values tended to be higher in black than

Table 3 R^2 and RMSE of regression analyses using %BF from body composition equations and %BF from DXA in the VIM I and Fels studies in Females

Equation	VIM I						Fels					
	Overall		Black		White		Overall		Adults		Youth	
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
B	---	---	---	---	---	---	**0.805**	3.07	**0.770**	3.06	**0.803**	2.99
C	---	---	---	---	---	---	**0.769**	3.35	0.734	3.29	0.744	3.41
D	---	---	---	---	---	---	**0.801**	3.11	**0.765**	3.09	**0.800**	3.01
E	---	---	---	---	---	---	0.745	3.51	0.711	3.42	0.707	3.65
F	---	---	---	---	---	---	**0.803**	3.09	**0.766**	3.08	**0.800**	3.01
G	**0.766**	4.25	0.682	4.00	**0.789**	4.34	**0.799**	3.12	**0.761**	3.12	**0.797**	3.04
H	---	---	---	---	---	---	**0.793**	3.17	**0.755**	3.15	**0.787**	3.11
I	**0.765**	4.26	0.687	3.98	**0.786**	4.37	**0.792**	3.17	**0.757**	3.14	**0.789**	3.09
J	0.691	4.89	0.579	4.61	0.721	4.98	0.741	3.54	0.704	3.47	0.710	3.63
K	0.685	4.94	0.582	4.59	0.711	5.08	0.722	3.67	0.685	3.58	0.674	3.85
L	0.686	4.93	0.582	4.60	0.713	5.06	0.732	3.60	0.696	3.51	0.701	3.68
M	**0.760**	4.31	0.681	4.02	**0.781**	4.42	**0.782**	3.25	0.746	3.21	**0.771**	3.22
N	0.677	5.00	0.576	5.17	0.700	5.17	0.713	3.73	0.677	3.62	0.659	3.93
1	---	---	---	---	0.699	5.18	---	---	0.643	3.80	---	---
2	---	---	---	---	0.740	4.81	---	---	0.624	3.90	---	---
3	---	---	---	---	0.635	5.70	---	---	0.583	4.11	---	---
4	0.511	6.15	0.400	5.51	0.564	6.23	---	---	0.613	3.96	---	---
5	---	---	---	---	0.747	4.75	---	---	0.681	3.60	---	---
6	---	---	---	---	0.717	5.02	---	---	0.662	3.70	---	---
7	---	---	---	---	0.729	4.92	---	---	0.699	3.49	---	---
8	---	---	---	---	0.737	4.84	---	---	0.685	3.58	---	---
9	---	---	---	---	---	---	---	---	---	---	0.740	3.54

R^2 greater than 0.750 are in bold. Variables in equations are shown in Table 1.

Table 4 R^2 and RMSE of regression analyses using %BF from body composition equations and %BF from DXA in the VIM I and Fels studies in Males

Equation	VIM I						Fels					
	Overall		Black		White		Overall		Adults		Youth	
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
B	---	---	---	---	---	---	**0.860**	2.43	**0.826**	2.44	**0.907**	2.34
C	---	---	---	---	---	---	**0.793**	2.96	**0.752**	2.91	**0.861**	2.87
D	---	---	---	---	---	---	**0.860**	2.44	**0.828**	2.42	**0.904**	2.39
E	---	---	---	---	---	---	**0.773**	3.10	0.737	3.00	**0.832**	3.16
F	---	---	---	---	---	---	**0.857**	2.46	**0.822**	2.47	**0.909**	2.33
G	0.706	3.70	0.747	3.68	0.625	3.60	**0.859**	2.44	**0.826**	2.44	**0.905**	2.37
H	---	---	---	---	---	---	**0.835**	2.64	**0.800**	2.61	**0.885**	2.61
I	0.702	3.73	**0.751**	3.65	0.604	3.70	**0.856**	2.47	**0.822**	2.46	**0.905**	2.37
J	0.539	4.63	0.594	4.66	0.424	4.46	**0.776**	3.08	0.739	2.98	**0.831**	3.16
K	0.486	4.89	0.533	5.00	0.381	4.62	0.672	3.72	0.635	3.53	0.735	3.96
L	0.536	4.65	0.600	4.63	0.415	4.49	**0.776**	3.08	0.740	2.98	**0.828**	3.19
M	0.642	4.08	0.688	4.09	0.533	4.01	**0.821**	2.75	**0.779**	2.75	**0.882**	2.65
N	0.491	4.87	0.546	4.93	0.362	4.69	0.661	3.79	0.619	3.61	0.729	4.01
1	---	---	---	---	0.475	4.25	---	---	0.620	3.60	---	---
2	---	---	---	---	0.527	4.04	---	---	0.703	3.19	---	---
3	---	---	---	---	0.390	4.59	---	---	0.723	3.08	---	---
4	0.542	4.62	0.621	4.50	0.369	4.66	---	---	0.693	3.24	---	---
5	---	---	---	---	0.519	4.07	---	---	0.686	3.28	---	---
6	---	---	---	---	0.437	4.41	---	---	0.624	3.58	---	---
7	---	---	---	---	0.551	3.94	---	---	0.732	3.02	---	---
8	---	---	---	---	0.538	3.99	---	---	0.693	3.24	---	---
9	---	---	---	---	---	---	---	---	---	---	**0.784**	3.60

R^2 greater than 0.750 are in bold. Variables in equations are shown in Table 1.

white men in the eight models studied. Without exception, the R^2 tended to be higher in Fels youth compared to adults for the same model.

Our examination of the MSD in subgroups was focused on detecting systematic differences in sub-groups. Figures 1 and 2 show that in the VIM I data the MSD in women, men, black and white participants was between −2 and 2 percentage points for all Stevens et al. equations except for K and N, which tended to overestimate in women. Equations 1, 2, 4, and 8 consistently underestimated %BF in the sex and ethnic subgroups studied. In the Fels data, Stevens et al. equations K, M, and N overestimated %BF in normal weight subjects. Estimates from Stevens et al. equations J and L were low in adults with obesity (MSD −2.07 and − 2.01 respectively). The other equations generally had a good to excellent MSD in normal weight adults, but %BF was underestimated by equations 1, 2, 4, 6, and 8. Equation 8 (Hassager) underestimated by the largest amount in normal weight and overweight adults while equation 3 (Kagawa) underestimated %BF the most in adults who had obesity. In the Fels data (Figures 3 and 4) the MSD estimates for the Stevens et al. equations were in the good to excellent range for all sex and ethnic groups and for the BMI status groups. None of the other equations were in that range for all the subgroups studied, however, the Durnin and Womersley (equation 1) and the Kagawa et al. (equation 3) MSD estimates were in the −2% and 2% percentage points body fat range for all except those with obesity.

Discussion

This was the first study to evaluate the external validity of the equations by Stevens and colleagues for the prediction of %BF, and results indicated that those equations generally have superior performance compared to several previously published equations that draw from the same list of anthropometric measurements (1). Given similar or equal performance, a smaller number of measurements in an equation is preferred since each measurement requires effort from the research team to collect and increases the duration of data collection when other assessments are also taking place potentially adding to the burden of study participants. Among the equations with high external validity, Stevens et al. equation I had

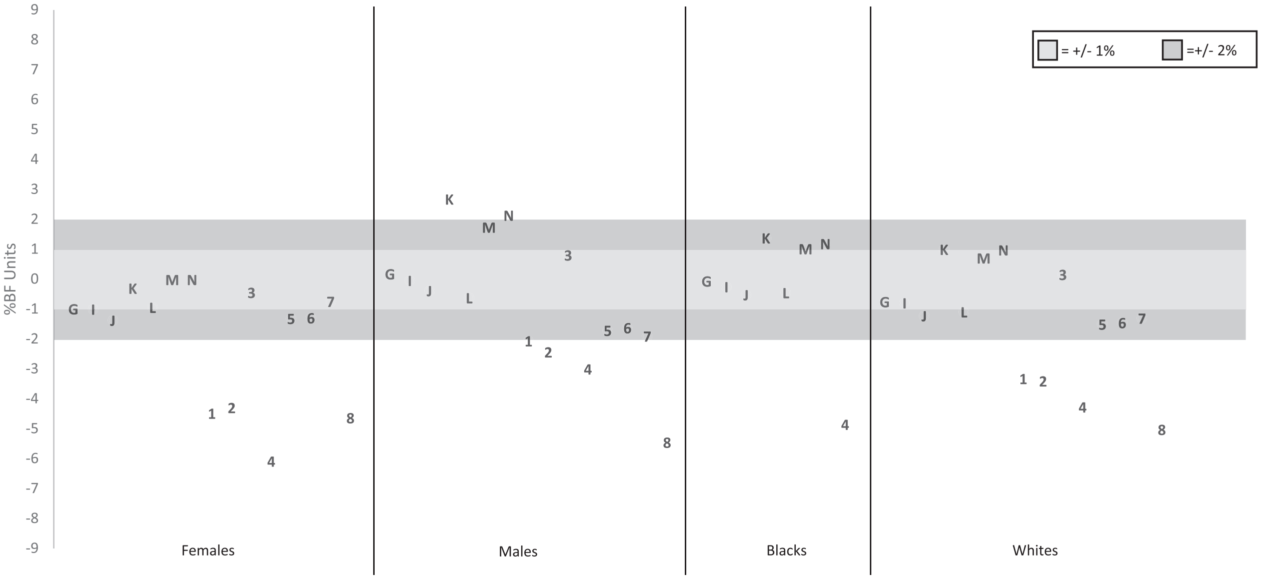

Figure 1 Mean signed differences in subgroups by sex and race categories in the VIM I sample. Letters and numbers indicate results from different equations (see Table 1 for equation identification). Values above 0 indicate overestimation and values below zero underestimation of percent body fat of equation estimates compared to DXA

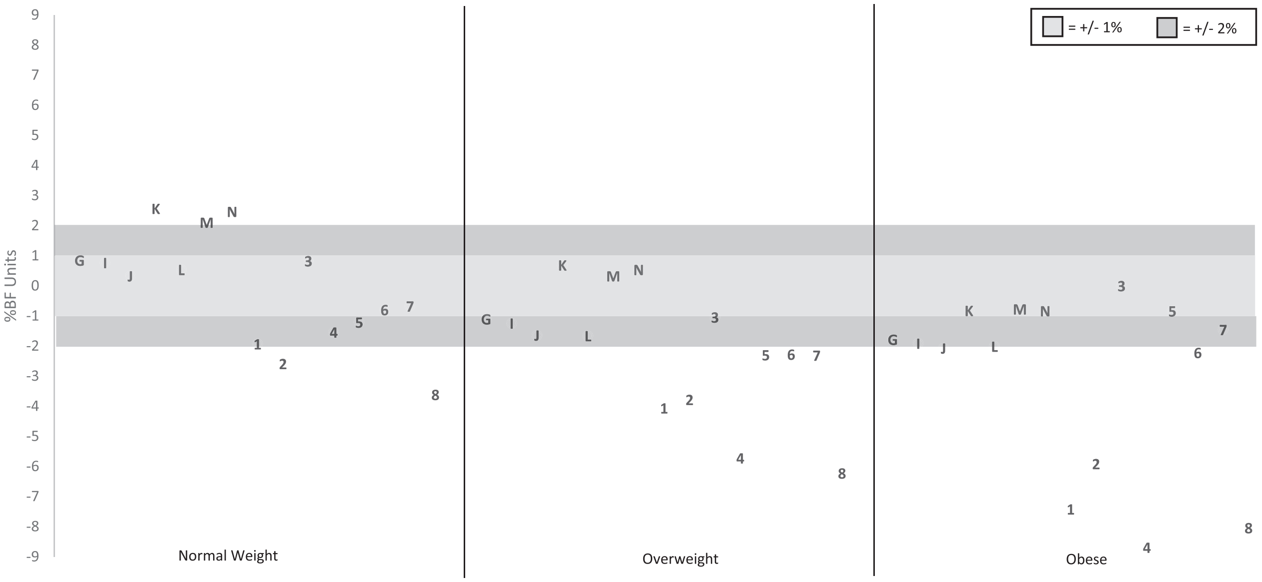

Figure 2 Mean signed differences in subgroups by BMI categories in the VIM I sample. Letters and numbers indicate results from different equations (see Table 1 for equation identification). Values above 0 indicate overestimation and values below zero underestimation of percent body fat of equation estimates compared to DXA.

the fewest variables, making it advantageous for use in future work. This equation consisted of the base variables (age, race, height, weight, BMI) plus triceps skinfold, and waist circumference.

There was a higher R^2 and lower RMSE values when the Stevens et al. equations were applied to the Fels study compared to the VIM I study. However, the MSD estimates were between −2 and 2 percentage points of body fat for almost all the Stevens et al. equations in both Fels and VIM I indicating no large difference in bias. As a frame of reference, a MSD of 2 percentage points of body fat corresponds to a percent error of 7% for a group with 30% body fat (the approximate overall mean observed in this study). Differences between protocols in the

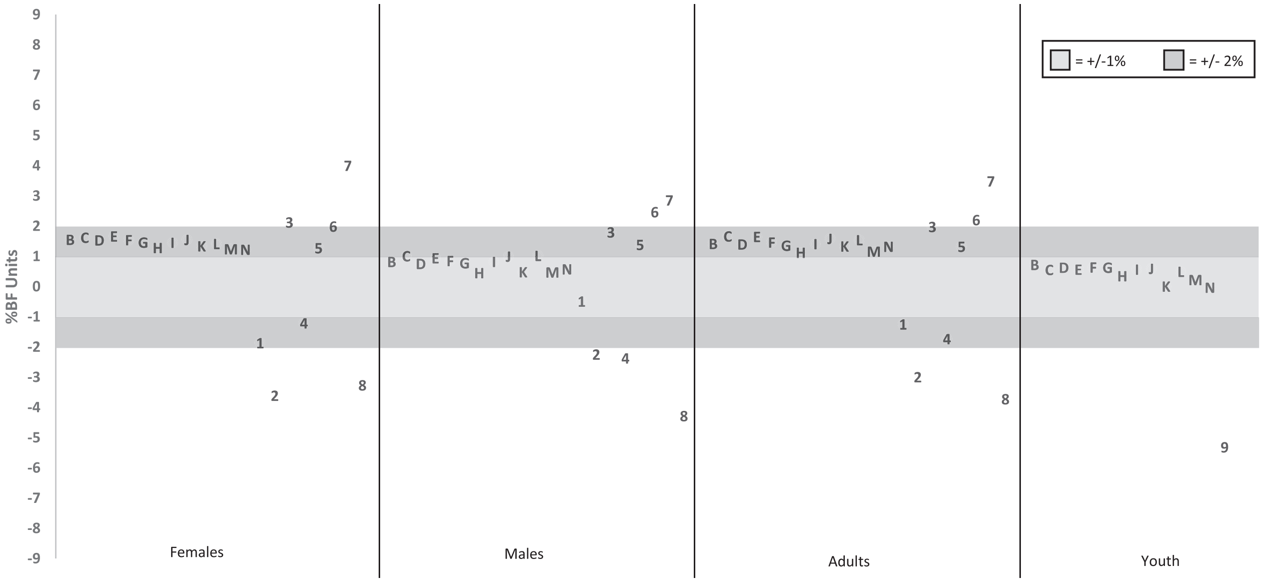

Figure 3 Mean signed differences in subgroups by sex and age categories in the Fels sample. Letters and numbers indicate results from different equations (see Table 1 for equation identification). Values above 0 indicate overestimation and values below zero underestimation of percent body fat of equation estimates compared to DXA.

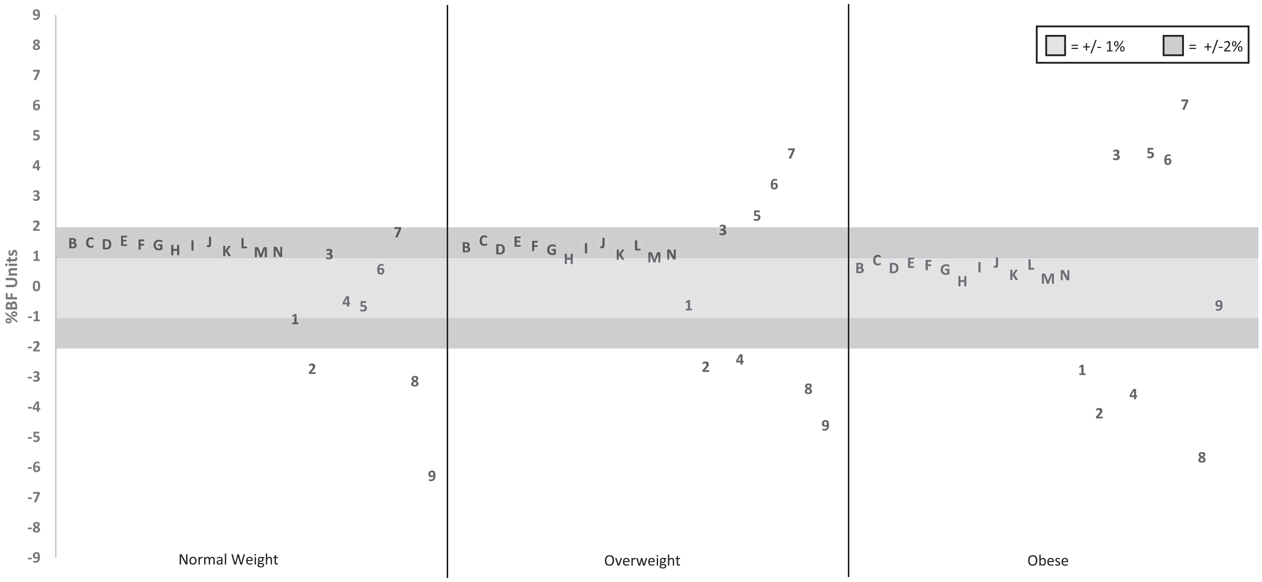

Figure 4 Mean signed differences in subgroups by BMI categories in the Fels sample. Letters and numbers indicate results from different equations (see Table 1 for equation identification). Values above 0 indicate overestimation and values below zero underestimation of percent body fat of equation estimates compared to DXA.

development and the external validation anthropometric methodology likely influenced the performance of the Stevens equations in VIM I as compared to Fels. In both Fels and NHANES measurement collection followed protocols described in the Anthropometric Standardization Reference Manual (8). In VIM I, a different anthropometric data collection protocol was used. Thus, it may be important for details of anthropometry methodology to be examined prior to application of equations for prediction of %BF.

Differences in criterion measures of %BF is a weakness that can also increase systematic and random error and result in reduced external validity. The Hologic QDR 4500A DXA was used as the criterion method in NHANES and Fels, whereas, the Hologic 2000 DXA was used as the criterion method in VIM I. A validation study conducted by Schoeller et al... showed that the Hologic QDR 4500A DXA underestimated body fat and overestimated lean soft tissue (5). A correction factor to improve accuracy was proposed by the authors, and this adjustment was applied to the NHANES and Fels data prior to its use in this study (5). There is no similar validation evidence or correction procedure for the Hologic 2000 that was used in VIM I. Underwater weighting (UWW), used as the criterion for equation development by Slaughter et al., Lean et al, and Durnin and Womersley, could have increased differences in comparisons made to the NHANES data collected using the Hologic QDR 4500A DXA. Similarly, the Dual Photon Absorptiometry (DPA) method, used by Hassager et al. for equation development, may have contributed to validation error.

In addition to differences in measurement methodology, the analytic methods used to develop equations differed across studies. Some investigators created and evaluated equation performance using the same data whereas others used separate data sets or random subsets of data from the same study for equation development versus internal validation (11,13) (10,15). Stevens et al. used different subsets of data for equation and validation, and studied variables in an unusually large number of mathematical forms (linear, squared, reciprocal and interactions). Models were selected from as many as 1,335 and 1,402 candidate terms for men and women, respectively. Evaluation of this large number of terms was made practical and stable by the use of the least absolute shrinkage and selection operator (LASSO) technique. Stevens and colleagues used the LASSO technique directed by the cross-validation error and adjusted R^2 to select terms and calculate model coefficients. The superior performance here of the Stevens et al. equations compared to the Zanovec et al. equations developed in the same data may have been due to the many variables and terms used in each Stevens et al. equation. Zanovec studied five variables in two equations and examined continuous anthropometric measures only in the linear form. Also, they did not conduct internal validation analyses.

Other studies have also examined the external validity of body composition estimates from equations. Davidson et al. showed differential bias in the estimate of %BF using Durnin and Womersley equations for white, black, Asian, and Hispanic men and women (16). The general and age specific Durnin and Womersley equations had the lowest mean error in black women and white men when compared to other race-sex subgroups. The mean error was higher when predicting %BF using the age-specific equations than the general equation for all subgroups (16). The Lean et al. equations developed in 16 to 65 year olds applied to youth 13 to 17 years resulted in estimates of %BF in men and women that were systematically low (17). Lean et al. and Durnin and Womersley equations underestimated %BF by 2.1% and 4.2% in non-disabled and by 10.6% and 8.3% in disabled athletes while Lean et al. equations overestimated %BF in post-menopausal women (18,19). Durnin and Womerley equations have been shown to underestimate %BF in men and women and to underestimate %BF more as BMI increases from overweight to obese (18,20–23) (17,23). Cui et al. showed that equations that included skinfolds tended to underestimate %BF more in the adults with obesity than the non-obese when applied to NHANES data (24). This trend was also observed in the present study among equations from the literature (MSD $>$ 2% BF) but not in the Stevens et al. equations.

Bias in subgroups may be due to under-representation in the equation development samples. The inclusion of participants with overweight or obesity may have been inadequate in the Slaughter et al. and Hassager et al. equation development samples in which the mean BMI was in the normal range, whereas Durnin and Womersley, Kagawa et al., and Lean et al. included larger proportions of participants with overweight and obesity in their equation development samples. Slaughter et al. equations have previously shown differential bias (non-zero differences in means) in men and women (21,25–27). In a sample of NHANES adults 20 to 29 years old the Slaughter et al. equation had an R^2 of 0.770 and 0.665 in men and women, respectively, with no bias in men but an underestimate of DXA %BF in women (24). Truesdale et al. examined the Slaughter et al. equations in NHANES youth and showed that they differentially underestimated percent body by 4.96 percentage points in boys and 6.48 percentage points in girls (28). In Pakistani children the Slaughter et al. equations had an R^2 of 0.76 and 75% of the estimates were within the limits of agreement (1 SD) between the DXA %BF and equation predicted body fat (i.e. ± 10 percentage points) (29). In contrast to the studies reviewed above, Wong et al. demonstrated that the Slaughter et al. equation exhibited small bias but a large standard error of the estimate in black and white women (30).

A limitation of the current study was inability to validate all the Stevens equations due to the absence of the necessary anthropometric variables in VIM I and Fels. Also, greater representation of ethnic groups, including Mexican Americans, in the both validation samples would have been desirable. The quality of the calculated estimates will likely be impacted by details of the procedures used to collected anthropometry, and is a limitation to the application of all the equations tested. In addition, all of the equations examined here may have notable levels of error when used for prediction at the individual level, and be more useful for the examination of groups. Finally, since both VMI I and Fels used DXA to produce criterion measures of percent body fat, the external validity assessments may have been attenuated for equations that used underwater weighing or DPA as the criterion. Nevertheless, this study has shown that the Stevens et al. equations are applicable in independent samples and generally provide superior estimates compared to several other equations in the literature.

Funding

This project is supported by R01-DK097046.

CARDIA is supported by contracts HHSN268201300025C, HHSN268201300026C, HHSN268201300027C, HHSN268201300028C, HHSN268201300029C, and HHSN268200900041C from the NHLBI, the IRP NIA, and an intra-agency agreement between NIA and NHLBI (AG0005).

FELS is supported by R01-HD-012252.

Disclosures

Ms. Reynolds has nothing to disclose.

Dr. Stevens has nothing to disclose.

Dr. Cai has nothing to disclose.

Dr. Lewis has nothing to disclose.

Dr. Czerwinski has nothing to disclose.

Dr. Choh has nothing to disclose.

References

1. Stevens J, Ou F-S, Cai J, Heymsfield SB, Truesdale KP. Prediction of percent body fat measurements in Americans 8 years and older. *International Journal of Obesity* 2016; **40**: 587–594. https://doi.org/10.1038/ijo.2015.231.
2. Roche A. *Growth, Maturation, and Body Composition: The Fels Longitudinal Study 1929–1991*, 1st edn. Cambridge University Press: Cambirdge, 1992.
3. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *Journal of Clinical Epidemiology* 1988; **41**: 1105–1116.
4. Hill JO, Sidney S, Lewis CE, Tolan K, Scherzinger AL, Stamm ER. Racial differences in amounts of visceral adipose tissue in young adults: the CARDIA (Coronary Artery Risk Development in Young Adults) study. *The American Journal of Clinical Nutrition* 1999; **69**: 381–387.
5. Schoeller DA, Tylavsky FA, Baer DJ, et al. QDR 4500A dual-energy X-ray absorptiometer underestimates fat mass in comparison with criterion methods in adults. *The American Journal of Clinical Nutrition* 2005; **81**: 1018–1025.
6. CARDIA Anthropometry Year 10 (1995–1996) Manual of Operation.
7. Fels Longitudinal Study Collection Methodologies.
8. Lohman TG, Roche AF, Martorell R. *Anthropometric Standardization Reference Manual*. Champaign, IL: Human Kinetics Publishers; 1988. http://search.lib.unc.edu?R=UNCb2182922.
9. Durnin JV, Womersley J. Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged from 16 to 72 years. *The British Journal of Nutrition* 1974; **32**: 77–97.
10. Lean ME, Han TS, Deurenberg P. Predicting body composition by densitometry from simple anthropometric measurements. *The American Journal of Clinical Nutrition* 1996; **63**: 4–14.
11. Zanovec M, Wang J, O'Neil CE. Development and comparison of two field-based body fat prediction equations: NHANES 1999-2004. *Int J Exerc Sci.* 2012; **5**: 223–231.
12. Hassager C, Gotfredsen A, Jensen J, Christiansen C. Prediction of body composition by age, height, weight, and skinfold thickness in normal adults. *Metabolism* 1986; **35**: 1081–1084.
13. Kagawa M, Byrne NM, Hills AP. Comparison of body fat estimation using waist:height ratio using different "waist" measurements in Australian adults. *The British Journal of Nutrition* 2008; **100**: 1135–1141. https://doi.org/10.1017/S0007114508966095.
14. Slaughter MH, Lohman TG, Boileau RA, et al. Skinfold equations for estimation of body fatness in children and youth. *Human Biology* 1988; **60**: 709–723.
15. Deurenberg P, Weststrate JA, Seidell JC. Body mass index as a measure of body fatness: age- and sex-specific prediction formulas. *The British Journal of Nutrition* 1991; **65**: 105–114.
16. Davidson LE, Wang J, Thornton JC, et al. Predicting fat percent by skinfolds in racial groups: Durnin and Womersley revisited. *Medicine and Science in Sports and Exercise* 2011; **43**: 542–549. https://doi.org/10.1249/MSS.0b013e3181ef3f07.
17. Rodriguez G, Moreno LA, Blay MG, et al. Body fat measurement in adolescents: comparison of skinfold thickness equations with dual-energy X-ray absorptiometry. *European Journal of Clinical Nutrition* 2005; **59**: 1158–1166. https://doi.org/10.1038/sj.ejcn.1602226.
18. Willems A, Paulson TAW, Keil M, Brooke-Wavell K, Goosey-Tolfrey VL. Dual-Energy X-Ray Absorptiometry, Skinfold Thickness, and Waist Circumference for Assessing Body Composition in Ambulant and Non-Ambulant Wheelchair Games Players. *Frontiers in Physiology* 2015; **6**: 356. https://doi.org/10.3389/fphys.2015.00356.
19. Kanellakis S, Manios Y. Validation of five simple models estimating body fat in white postmenopausal women: use in clinical practice and research. *Obesity (Silver Spring)* 2012; **20**: 1329–1332. https://doi.org/10.1038/oby.2011.403.
20. Visser M, van den Heuvel E, Deurenberg P. Prediction equations for the estimation of body composition in the elderly using anthropometric data. *The British Journal of Nutrition* 1994; **71**: 823–833.
21. Bandara T, Hettiarachchi M, Liyanage C, Amarasena S, Wong WW-L. Body composition among Sri Lankan infants by (1)(8) O dilution method and the validity of anthropometric equations to predict body fat against (1)(8) O dilution. *BMC Pediatrics* 2015; **15**: 52. https://doi.org/10.1186/s12887-015-0371-2.
22. Peterson MJ, Czerwinski SA, Siervogel RM. Development and validation of skinfold-thickness prediction equations with a 4-compartment model. *The American Journal of Clinical Nutrition* 2003; **77**: 1186–1191.
23. Dioum A, Gartner A, Maire B, Delpeuch F, Wade S. Body composition predicted from skinfolds in African women: a cross-validation study using air-displacement plethysmography and a black-specific equation. *The British Journal of Nutrition* 2005; **93**: 973–979.

24. Cui Z, Truesdale KP, Cai J, Stevens J. Evaluation of anthropometric equations to assess body fat in adults: NHANES 1999-2004. *Medicine and Science in Sports and Exercise* 2014; **46**: 1147–1158. https://doi.org/10.1249/MSS.0000000000000213.
25. Loftin M, Nichols J, Going S, et al. Comparison of the validity of anthropometric and bioelectric impedance equations to assess body composition in adolescent girls. *Int J Body Compos Res.* 2007; **5**: 1–8.
26. Noradilah MJ, Ang YN, Kamaruddin NA, Deurenberg P, Ismail MN, Poh BK. Assessing Body Fat of Children by Skinfold Thickness, Bioelectrical Impedance Analysis, and Dual-Energy X-Ray Absorptiometry: A Validation Study Among Malay Children Aged 7 to 11 Years. *Asia-Pacific J public Heal.* 2016; **28**: 74S–84S. https://doi.org/10.1177/1010539516641505.
27. Reilly JJ, Wilson J, Durnin JV. Determination of body composition from skinfold thickness: a validation study. *Archives of Disease in Childhood* 1995; **73**: 305–310.
28. Truesdale KP, Roberts A, Cai J, Berge JM, Stevens J. Comparison of Eight Equations That Predict Percent Body Fat Using Skinfolds in American Youth. *Child Obes.* 2016; **12**: 314–323. https://doi.org/10.1089/chi.2015.0020.
29. Hussain Z, Jafar T, Zaman MU, Parveen R, Saeed F. Correlations of skin fold thickness and validation of prediction equations using DEXA as the gold standard for estimation of body fat composition in Pakistani children. *BMJ Open* 2014; **4**: e004194. https://doi.org/10.1136/bmjopen-2013-004194.
30. Wong WW, Stuff JE, Butte NF, Smith EO, Ellis KJ. Estimating body fat in African American and white adolescent girls: a comparison of skinfold-thickness equations with a 4-compartment criterion model. *The American Journal of Clinical Nutrition* 2000; **72**: 348–354.