MAYO
CLINIC

Do Missing Values Influence Outcomes in a Cross-sectional Mail Survey?

Paul J. Novotny, MS; Darrell Schroeder, MS; Jeff A. Sloan, PhD; Gina L. Mazza, PhD; David Williams, PhD; David Bradley, MD; Irina V. Haller, PhD; Steven M. Bradley, MD, MPH; and Ivana Croghan, MS, PhD

Abstract

Objective: To determine the effects of missing and inconsistent data on a weight management mail survey results.

Patients and Methods: Weight management surveys were sent to 5000 overweight and obese individuals in the Learning Health System Network. Survey information was collected between October 27, 2017, and March 1, 2018. Some participants reported body mass index (BMI) values inconsistent with the intended overweight and obese sampling cohort. Analyses were performed after excluding these surveys and also performed again after setting these low BMI values to missing. Models were run after imputing missing values using expectation-maximization, Markov chain Monte Carlo, random forest imputation, multivariate imputation by chained equations, and multiple imputation and replacing missing BMI values with the minimum, maximum, mean, or median of the known BMI values.

Results: Of 2799 surveys, 222 (8%) had missing BMI values and 155 (6%) reported invalid BMI values. Overall, 725 of these 2799 surveys (26%) were missing at least 1 variable that was essential to the main analyses. Different imputation methods consistently found that BMI was related to age, sex, race, marital status, and education. Patients with a BMI of 35.0 kg/m^2 or greater were more likely to feel judged because of their weight, and patients with a BMI of 40.0 kg/m^2 or greater were more likely to feel they were not always treated with respect and treated as an equal.

Conclusion: Analyses using different imputation methods were consistent with the original published results. Missing data likely did not affect the study results.

From the Division of Biomedical Statistics and Informatics (P.J.N., D.S., J.A.S.) and Department of Medicine (I.C.), Mayo Clinic, Rochester, MN; Division of Biomedical Statistics and Informatics, Mayo Clinic, Scottsdale, AZ (G.L.M.); Department of Anesthesiology, University of Michigan, Ann Arbor (D.W.); Diabetes and Metabolism Research Center, Division of Endocrinology, Diabetes and Metabolism, Department of Internal Medicine, The Ohio State University, Columbus (D.B.); Essentia Institute of

Affiliations continued at the end of this article.

Missing data are inevitable in mail surveys. Missing data are rarely missing completely at random (MCAR) and can be due to patient address changes, death, inability (too ill) to respond, or refusal to answer the survey. Patients who are older and more frail with more hectic daily lives and more chronic conditions are less likely to respond.[1] In addition to reducing the statistical power of the study, missing data can lead to bias in the study results, skewing estimates away from the true parameter values the investigators are trying to measure. The usual method of handling missing data is to remove observations that have missing data for any of the variables used in the analysis and report results that ignore the missing data.[2] Rough guidelines for this approach suggest that if less than 10% of study participants have missing data for 1 or more of the analysis variables, then the study results should not be greatly affected. If more than 40% of observations have missing values, then variables with the most missing values should be removed from the analyses.[3]

The aim of the present project was to investigate the effects of missing data on the analysis results of a cross-sectional survey of overweight and obese patients.[4] The objective of the survey was to assess weight management needs of overweight and obese patients within the Learning Health System Network. The Learning Health System Network[4] is a research collaboration setup to facilitate cooperative research.

PATIENTS AND METHODS

This is a reanalysis of the data from a weight management survey of patients in 5 sites in the Learning Health System Network. At each participating site, surveys were sent to 1000 randomly selected patients in each of the 4 strata defined according to body mass index (BMI, calculated as the weight in kilograms divided by the height in meters squared): overweight, 25.0 to 25.9 kg/m^2; obesity class I, 30.0 to 34.9 kg/m^2; obesity class II, 35.0 to 39.9 kg/m^2; and obesity class III, 40.0 kg/m^2 and greater. Survey information was collected between October 27, 2017, and March 1, 2018. Details of the data collection protocol are provided in the original article.[4] Primary analyses assessed the association between obesity and respondents' perceptions of their primary care provider's behavior. Three respondent perceptions of their primary care provider were of specific interest: "being judged because of your weight," "not always treated with respect," and "not always treated as an equal." Each of these perceptions was measured using a binary variable and analyzed using multiple logistic regression. In addition to BMI category (25.0-25.9 kg/m^2 vs 30.0-34.9 kg/m^2 vs 35.0-39.9 kg/m^2 vs ≥40.0 kg/m^2), the other covariates included in the models were age (treated as a continuous variable), sex (male vs female), race (non-Hispanic white vs other), marital status (married/living as married vs other), education (high school graduate or less vs some college vs 4-year college degree or more), and the presence of multiple comorbidities (yes vs no).

In the original report,[4] respondents were excluded from all primary analyses if their reported BMI fell below the lower limit used for study inclusion (ie, BMI <25.0 kg/m^2) or if their age, sex, or BMI were missing. In addition to these exclusions, the analysis data set had sporadic missing data for other covariates and end points included in the logistic regression models. In the original report, these respondents were also excluded from the analyses. In the present study, analyses from the original report are repeated using a myriad of different methods for handling missing data.

This study was overseen by the Mayo Clinic Institutional Review Board, which determined that the study was exempt under 45 CFR 46.101, item 2. Protocol-approved passive consent was obtained from all study participants.

Statistical Methods

The survey aimed to assess only those individuals who were overweight. Because the survey was intended to be completed only by patients with a BMI of 25.0 kg/m^2 or greater, respondents were first categorized into 3 groups: missing BMI, BMI less than 25.0 kg/m^2, and BMI 25.0 kg/m^2 or greater. To assess characteristics associated with missing or potentially incorrect BMI information, differences in demographic characteristics between these groups were tested using chi-square tests for categorical variables and Kruskal-Wallis nonparametric tests for continuous variables. Similarly, the association between missingness of each sociodemographic variable (0=reported; 1=missing) and the observed values of each of the remaining sociodemographic variables was assessed using logistic regression.

To assess the potential effect of missing item-level data on the results presented in the original report, 11 methods were used for handling the missing data: (1) excluding respondents with missing data, (2) imputing the minimum BMI for any missing BMI values, (3) imputing the maximum BMI for any missing BMI values, (4) imputing the mean BMI for any missing BMI values, (5) imputing the median BMI for any missing BMI values, (6) using the expectation-maximization algorithm to impute BMI and all other incomplete variables, (7) using the Markov chain Monte Carlo (MCMC) method to impute BMI and all other incomplete variables, (8) using random forest imputation (using R package missForest which is free software available through the Comprehensive R Archive Network) to impute BMI and all other incomplete variables, (9) multiple imputation while creating 10 imputed data sets,[5] (10) multivariate imputation by chained equations, and (11) tipping point multiple imputation sensitivity analysis. For multiple imputation, the MCMC method is first used to fill in just enough missing data to make the data monotone. This means that the variables are in order

from the lowest number of missing values, and if a variable is missing, then all variables after that variable are also missing. After the data are filled in enough to be monotone, each variable is then imputed 1 at a time in order from the variable with the lowest proportion of missing data to the variable with the highest proportion of missing data. Separate regression models (logistic regression models for binary variables and linear regression models for continuous variables) are used for each variable to sequentially impute the missing values on the basis of the variables that have been imputed or are not missing.

Method 1 (excluding those with item-level missing data) is the approach used in the original analysis. Methods 2 to 5 use a single summary statistic to impute a BMI value for all respondents with missing BMI. Each of these approaches results in a single analysis data set in which all respondents have data for BMI, but missing data for other covariates are not imputed. Therefore, respondents who have missing data for other covariates will be excluded from the analyses. Methods 6 to 8 impute plausible values for all variables with missing data by taking into account the correlation structure between the reported values of BMI, age, sex, race, marital status, presence of multiple comorbidities, education, and outcome variables. These methods were used to create a single imputed analysis data set that has complete data for all variables. Method 9 also uses the correlation structure of the observed data, but instead of imputing a single plausible value for each missing data point, 10 plausible values are imputed. This results in 10 imputed analysis data sets each with complete data for all variables. For this approach, analyses were performed separately for each imputed data set, with the results combined using Rubin's rules.[6] This multiple imputation method adjusts the SEs to account for uncertainty due to the missing scores.

Multivariate imputation by chained equations is a recursive method of filling in missing values for all the variables in the data set. First, the missing values are all filled in with rough estimates such as the mean values. Then, the missing values for one of the variables are reset to missing and are estimated using regression estimates on the basis of the values of all the other variables by using both the real and imputed values. Next, the missing values of another variable are reset to missing and estimated using the real and imputed values of the other variables. This process is repeated for all the variables and then the entire process is redone until the imputed values no longer change substantially.

All the imputation methods (1-10) are reasonable methods if the missing values are missing at random (MAR) or MCAR. Missing completely at random means that the missing values are not related to any observed or unobserved variables. Missing at random means the missing data are a function of variables that are observed in the data set. However, it is almost always the case that missing values are missing for an unknown reason. These values are considered missing not at random (MNAR). Although they could be missing for an almost infinite number of reasons, it is possible to model the MNAR mechanism to determine how sensitive the results are to different MNAR models.[7,8] We used the MNAR option in SAS proc MI (SAS Institute Inc, Cary, North Carolina, USA) to assess the effects of changing the log odds of a response in individuals with missing values. These models were used to find tipping points at which the results of the original analyses changed.

After creating an analysis data set for each missing data approach, multiple logistic regression analyses were performed to assess the association of obesity with respondent perceptions of their primary care provider's behavior. As in the original analyses, age, sex, race, marital status, education, and presence of multiple comorbidities were included as covariates. The results from these analyses are summarized by presenting the odds ratio (OR) and 95% CI for each of the BMI categories, using those with a BMI of 25.0 to 25.9 kg/m^2 as the reference.

The missing data analyses used a macro concurrently being developed by Dr Jeff A. Sloan, Dr Amylou Dueck, and Mr Paul J. Novotny at Mayo Clinic.[9] This macro combines SAS and R code to analyze patient-reported outcomes with multiple time points. But in this case, we found an additional use for this macro in a cross-sectional survey with only 1 time point.

Characteristic	Missing BMI (n=222)	BMI <25.0 kg/m² (n=155)	BMI 25.0-29.9 kg/m² (n=703)	BMI 30.0-34.9 kg/m² (n=665)	BMI 35.0-39.9 kg/m² (n=503)	BMI ≥40.0 kg/m² (n=551)
TABLE 1. Demographic Characteristics by BMI Group[a,b]						
Reported BMI						
n	0	155	703	665	503	551
Mean (kg/m2)		23.8	27.5	32.4	37.3	46.1
Age						
n	211	150	685	656	498	542
Mean (y)	65.7	61.0	62.1	60.9	58.8	54.3
Sex						
Missing	8 (4)	2 (1)	17 (2)	7 (1)	4 (1)	8 (1)
Female	128 (58)	88 (57)	360 (51)	359 (54)	321 (64)	404 (73)
Male	86 (39)	65 (42)	326 (46)	299 (45)	178 (35)	138 (25)
Other	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0)
Race						
Missing	11 (5)	2 (1)	22 (3)	9 (1)	6 (1)	8 (1)
Asian	0 (0)	4 (3)	8 (1)	1 (0)	0 (0)	1 (0)
Black	14 (6)	2 (1)	14 (2)	22 (3)	14 (3)	28 (5)
Other	6 (3)	9 (6)	21 (3)	18 (3)	24 (5)	29 (5)
White	191 (86)	138 (89)	638 (91)	615 (92)	459 (91)	485 (88)
Marital status						
Missing	120 (54)	2 (1)	7 (1)	6 (1)	10 (2)	10 (2)
Married	63 (28)	117 (75)	521 (74)	497 (75)	358 (71)	330 (60)
Never married	7 (3)	14 (9)	47 (7)	48 (7)	41 (8)	89 (16)
Separated/divorced	11 (5)	8 (5)	75 (11)	74 (11)	58 (12)	89 (16)
Widowed	21 (9)	14 (9)	53 (8)	40 (6)	36 (7)	33 (6)
Education						
Missing	125 (56)	2 (1)	10 (1)	12 (2)	10 (2)	12 (2)
Less than HS graduate	6 (3)	5 (3)	18 (3)	16 (2)	13 (3)	9 (2)
HS graduate	28 (13)	25 (16)	117 (17)	117 (18)	91 (18)	115 (21)
Some college	32 (14)	35 (23)	223 (32)	224 (34)	183 (36)	239 (43)
4-Y college degree	10 (5)	29 (19)	151 (21)	130 (20)	104 (21)	86 (16)
Some postgraduate	6 (3)	9 (6)	34 (5)	37 (6)	22 (4)	25 (5)
Postgraduate or professional degree	15 (7)	50 (32)	150 (21)	129 (19)	80 (16)	65 (12)
Multiple comorbidities						
No	158 (71)	91 (59)	328 (47)	211 (32)	145 (29)	124 (23)
Yes	64 (29)	64 (41)	375 (53)	454 (68)	358 (71)	427 (77)
Judged because of weight						
Missing	32 (14)	9 (6)	50 (7)	57 (9)	48 (10)	69 (13)
No	174 (78)	141 (91)	638 (91)	577 (87)	407 (81)	417 (76)
Yes	16 (7)	5 (3)	15 (2)	31 (5)	48 (10)	65 (12)
Not always treated with respect						
Missing	20 (9)	6 (4)	35 (5)	22 (3)	17 (3)	19 (3)
No	163 (73)	134 (86)	586 (83)	548 (82)	411 (82)	422 (77)
Yes	39 (18)	15 (10)	82 (12)	95 (14)	75 (15)	110 (20)
Not always treated as an equal						
Missing	19 (9)	6 (4)	40 (6)	26 (4)	20 (4)	23 (4)
No	152 (68)	111 (72)	506 (72)	476 (72)	353 (70)	354 (64)
Yes	51 (23)	38 (25)	157 (22)	163 (25)	130 (26)	174 (32)

Continued on next page

TABLE 1. Continued						
Characteristic	Missing BMI (n=222)	BMI <25.0 kg/m^2 (n=155)	BMI 25.0-29.9 kg/m^2 (n=703)	BMI 30.0-34.9 kg/m^2 (n=665)	BMI 35.0-39.9 kg/m^2 (n=503)	BMI ≥40.0 kg/m^2 (n=551)
General health						
Missing	120 (54)	0 (0)	3 (0)	8 (1)	2 (0)	5 (1)
Excellent	5 (2)	38 (25)	64 (9)	23 (3)	3 (1)	6 (1)
Very good	29 (13)	62 (40)	285 (41)	195 (29)	117 (23)	65 (12)
Good	43 (19)	43 (28)	266 (38)	295 (44)	246 (49)	235 (43)
Fair	19 (9)	9 (6)	73 (10)	123 (18)	112 (22)	193 (35)
Poor	6 (3)	3 (2)	12 (2)	21 (3)	23 (5)	47 (9)
Positive screen result for current depression (PHQ-2)						
Missing	30 (14)	10 (6)	61 (9)	68 (10)	37 (7)	40 (7)
No	156 (70)	140 (90)	584 (83)	513 (77)	392 (78)	363 (66)
Yes	36 (16)	5 (3)	58 (8)	84 (13)	74 (15)	148 (27)
Currently smoke cigarettes						
Missing	119 (54)	1 (1)	3 (0)	2 (0)	3 (1)	1 (0)
Yes	9 (4)	9 (6)	39 (6)	44 (7)	26 (5)	36 (7)
No	94 (42)	145 (94)	661 (94)	619 (93)	474 (94)	514 (93)
Currently use alcohol products						
Missing	125 (56)	9 (6)	26 (4)	24 (4)	22 (4)	10 (2)
Yes	41 (18)	86 (55)	458 (65)	381 (57)	269 (53)	251 (46)
No	56 (25)	60 (39)	219 (31)	260 (39)	212 (42)	290 (53)
Considered overweight as a child						
Missing	123 (55)	0 (0)	3 (0)	6 (1)	3 (1)	7 (1)
Yes	19 (9)	20 (13)	92 (13)	141 (21)	177 (35)	268 (49)
No	80 (36)	135 (87)	608 (86)	518 (78)	323 (64)	276 (50)
Current opinion of their weight						
Missing	28 (13)	11 (7)	63 (9)	71 (11)	43 (9)	44 (8)
Underweight	1 (0)	0 (0)	1 (0)	2 (0)	1 (0)	0 (0)
Average or normal weight	45 (20)	111 (72)	249 (35)	40 (6)	4 (1)	2 (0)
Overweight	95 (43)	31 (20)	376 (53)	429 (65)	214 (43)	100 (18)
Obese	42 (19)	2 (1)	13 (2)	118 (18)	207 (41)	204 (37)
Very obese	11 (5)	0 (0)	1 (0)	5 (1)	34 (7)	201 (36)
Physical violence with growing up						
Missing	25 (11)	9 (6)	60 (9)	61 (9)	32 (6)	34 (6)
Yes	28 (13)	17 (11)	98 (14)	103 (15)	104 (21)	129 (23)
No	169 (76)	129 (83)	545 (78)	501 (75)	367 (73)	388 (70)

[a]BMI = body mass index; HS = high school; PHQ-2 = Patient Health Questionnaire-2.
[b]Data are expressed as No. (percentage) unless indicated otherwise.

The survey was intended to be sent only to patients with a BMI of 25.0 kg/m^2 or greater. Therefore, respondents who reported a BMI of less than 25.0 kg/m^2 are problematic. It cannot be determined with certainty whether their BMI levels have actually decreased or whether they reported lower BMI levels because lower values are more socially desirable. So it is not clear whether their survey results should be retained in our survey of overweight individuals. Because no patient identifying information was included in the returned surveys, it was impossible to use medical record information to assess the potential accuracy/validity of these values. For the original analysis, these respondents were excluded. For the main analyses presented in the present study, these respondents are also excluded. To assess the sensitivity of the results

TABLE 2. Extent of Missing Data[a,b]

Cohort	All surveys	Excluding surveys with BMI <25.0 kg/m^2
Total n	2799	2644
Missing BMI	377 (14)	222 (8)
Missing judged because of weight	265 (10)	256 (10)
Missing always treated with respect	119 (4)	113 (4)
Missing always treated as an equal	134 (5)	128 (5)
Missing age	57 (2)	52 (2)
Missing sex	46 (2)	44 (2)
Missing race	58 (2)	56 (2)
Missing marital status	155 (6)	153 (6)
Missing education	171 (6)	169 (6)
Missing multiple comorbidities	0 (0)	0 (0)
Missing any of these variables	725 (26)	570 (22)

[a]BMI = body mass index.
[b]Data are expressed as No. (percentage) unless indicated otherwise.

to these exclusions, we repeated all analyses from the present report while using data from all respondents, with BMI set to missing if the reported value was less than 25.0 kg/m^2. The results of these additional analyses are presented in Supplemental Tables 1 to 4 (available online at http://www.mcpiqojournal.org).

All analyses were 2-sided with 5% type I error rates. No adjustments were done for multiple testing. Analyses were performed using both R version 3.4.2 (Comprehensive R Archive Network) and SAS version 9.4 (SAS Institute Inc, Cary, North Carolina, USA).

RESULTS

Patient Demographic Characteristics

Of 19,964 mailed surveys, 15,819 (79%) had no response, 313 (2%) were returned by the

TABLE 3. P Values for Associations With BMI Using Different Imputation Methods Excluding Participants With a BMI of <25.0 kg/m^2[a,b]

Imputation method	Age: P value	Female: P value	non-Hispanic white: P value	Single: P value	High school education: P value	Some college: P value
Original results	<.001	<.001	.018	<.001	.49	<.001
Minimum	<.001	<.001	.028	<.001	.74	<.001
Maximum	<.001	<.001	.033	<.001	.09	<.001
Mean	<.001	<.001	.022	<.001	.44	<.001
Median	<.001	<.001	.022	<.001	.44	<.001
EM algorithm	<.001	<.001	.005	<.001	.29	<.001
MCMC algorithm	<.001	<.001	.008	<.001	.34	<.001
Random forest imputation	<.001	<.001	<.001	<.001	.47	<.001
MICE imputation	<.001	<.001	<.001	<.001	.07	<.001

[a]P values are based on univariate logistic regression models.
[b]BMI = body mass index; EM = expectation-maximization; MCMC = Markov chain Monte Carlo; MICE = multivariate imputation by chained equations.

TABLE 4. Logistic Model for Feeling Judged by BMI Excluding Patients With Low BMI[a,b]

Imputation	Type III: P value	BMI 30.0-34.9 kg/m² : odds ratio (95% CI)	P value	BMI 35.0-39.9 kg/m² : odds ratio (95% CI)	P value	BMI ≥40.0 kg/m² : odds ratio (95% CI)	P value
Original results	<.001	2.38 (1.22-4.63)	.011	4.62 (2.45-8.74)	<.001	5.26 (2.78-9.96)	<.001
Minimum	<.001	2.13 (1.14-3.98)	.017	4.14 (2.29-7.48)	<.001	4.68 (2.59-8.46)	<.001
Maximum	<.001	2.35 (1.21-4.58)	.012	4.55 (2.41-8.61)	<.001	4.76 (2.54-8.95)	<.001
Mean	<.001	2.34 (1.21-4.51)	.011	4.59 (2.43-8.69)	<.001	5.20 (2.75-9.84)	<.001
Median	<.001	2.34 (1.21-4.51)	.011	4.59 (2.43-8.69)	<.001	5.20 (2.75-9.84)	<.001
EM algorithm	<.001	2.38 (1.22-4.61)	.011	4.60 (2.44-8.67)	<.001	5.18 (2.74-9.82)	<.001
MCMC algorithm	<.001	2.35 (1.21-4.56)	.012	4.54 (2.40-8.57)	<.001	5.28 (2.79-10.00)	<.001
Random forest imputation	<.001	2.32 (1.20-4.50)	.013	4.51 (2.39-8.50)	<.001	5.28 (2.79-9.99)	<.001
Multiple imputation	<.001	2.00 (1.07-3.75)	.030	3.82 (2.07-7.03)	<.001	4.70 (2.62-8.44)	<.001
MICE	<.001	2.32 (1.22-4.39)	.010	4.30 (2.39-7.74)	<.001	6.12 (3.35-11.18)	<.001

[a]BMI = body mass index; EM = expectation-maximization; MCMC = Markov chain Monte Carlo; MICE = multivariate imputation by chained equations.
[b]Models were adjusted for age, sex, race, marital status, education, and presence of multiple comorbidities as covariates.

post office because of invalid addresses, and 1033 (5%) were returned by patients but did not include responses to any of the questions. This resulted in 2799 completed or partially completed surveys, which represented 14% of the mailed surveys.

Of the 2799 completed surveys, 2422 (87%) reported height and weight consistent with a BMI of 25.0 kg/m² or greater, 222 (8%) were missing BMI values, and another 155 (6%) reported BMI values that were inconsistent with being overweight or obese (ie, <25.0 kg/m²). Table 1 presents demographic characteristics for participants within categories of reported BMI. Participants with missing BMI values were older, more likely to be black, less likely to be married/living as married, less likely to have a 4-year college degree, less likely to have multiple comorbidities, less likely to rate their health as excellent, and less likely to use alcohol products than participants with observed BMI values.

TABLE 5. Logistic Model for Not Always Treated With Respect by BMI Excluding Patients With Low BMI[a,b]

Imputation	Type III: P value	BMI 30.0-34.9 kg/m² : odds ratio (95% CI)	P value	BMI 35.0-39.9 kg/m² : odds ratio (95% CI)	P value	BMI ≥40.0 kg/m² : odds ratio (95% CI)	P value
Original results	.11	1.24 (0.89-1.74)	.20	1.10 (0.76-1.57)	.62	1.51 (1.07-2.14)	.021
Minimum	.13	1.20 (0.87-1.66)	.26	1.06 (0.75-1.51)	.74	1.46 (1.04-2.04)	.028
Maximum	.13	1.24 (0.89-1.72)	.21	1.09 (0.76-1.56)	.65	1.46 (1.04-2.04)	.027
Mean	.11	1.24 (0.90-1.71)	.20	1.09 (0.76-1.57)	.64	1.50 (1.06-2.12)	.023
Median	.11	1.24 (0.90-1.71)	.20	1.09 (0.76-1.57)	.64	1.50 (1.06-2.12)	.023
EM algorithm	.12	1.27 (0.91-1.76)	.15	1.14 (0.80-1.63)	.47	1.51 (1.06-2.13)	.021
MCMC algorithm	.07	1.27 (0.92-1.76)	.15	1.08 (0.76-1.55)	.66	1.53 (1.08-2.17)	.016
Random forest imputation	.10	1.22 (0.88-1.69)	.24	1.07 (0.75-1.53)	.70	1.50 (1.06-2.12)	.022
Multiple imputation	.25	1.19 (0.86-1.64)	.29	1.19 (0.85-1.66)	.31	1.41 (1.00-1.99)	.048
MICE	.07	1.18 (0.84-1.66)	.33	1.15 (0.81-1.64)	.42	1.42 (1.02-1.98)	.037

[a]BMI = body mass index; EM = expectation-maximization; MCMC = Markov chain Monte Carlo; MICE = multivariate imputation by chained equations.
[b]Models were adjusted for age, sex, race, marital status, education, and presence of multiple comorbidities as covariates.

Imputation	Type III: P value	BMI 30.0-34.9 kg/m^2: odds ratio (95% CI)	P value	BMI 35.0-39.9 kg/m^2: odds ratio (95% CI)	P value	BMI ≥40.0 kg/m^2: odds ratio (95% CI)	P value
Original results	.12	1.06 (0.81-1.38)	.68	1.05 (0.79-1.40)	.74	1.37 (1.03-1.82)	.030
Minimum	.12	1.07 (0.83-1.39)	.58	1.06 (0.81-1.41)	.66	1.38 (1.05-1.82)	.022
Maximum	.34	1.05 (0.80-1.36)	.73	1.03 (0.78-1.37)	.82	1.26 (0.96-1.65)	.100
Mean	.13	1.02 (0.79-1.32)	.87	1.04 (0.78-1.38)	.79	1.35 (1.01-1.79)	.039
Median	.13	1.02 (0.79-1.32)	.87	1.04 (0.78-1.38)	.79	1.35 (1.01-1.79)	.039
EM algorithm	.14	1.04 (0.80-1.35)	.76	1.05 (0.80-1.39)	.72	1.35 (1.02-1.79)	.036
MCMC algorithm	.10	1.04 (0.80-1.35)	.76	1.05 (0.79-1.39)	.74	1.38 (1.04-1.83)	.025
Random forest imputation	.07	1.02 (0.78-1.32)	.90	1.01 (0.76-1.34)	.93	1.38 (1.04-1.82)	.027
Multiple imputation	.16	1.08 (0.85-1.38)	.52	1.12 (0.86-1.46)	.41	1.35 (1.03-1.77)	.029
MICE	.06	1.11 (0.86-1.44)	.43	1.14 (0.87-1.50)	.35	1.41 (1.08-1.85)	.012

TABLE 6. Logistic Model for Not Always Treated as an Equal by BMI Excluding Patients With Low BMI[a,b]

[a]BMI = body mass index; EM = expectation-maximization; MCMC = Markov chain Monte Carlo; MICE = multivariate imputation by chained equations.
[b]Models were adjusted for age, sex, race, marital status, education, and presence of multiple comorbidities as covariates.

Participants with BMI less than 25.0 kg/m^2 (not overweight or obese) were less likely to have a positive depression screen result and were more likely to consider themselves average or normal weight than other participants. Participants reporting BMI consistent with being overweight or obese were more likely to be considered overweight as a child and were more likely to have experienced physical violence when growing up than participants with missing BMI or BMI less than 25.0 kg/m^2.

Overall, 725 of 2799 surveys (26%) were missing at least 1 of the covariates included in the logistic regression analyses (Table 2).

The mean BMI of 35.1 kg/m^2 from the returned surveys was significantly ($P<.001$) lower than the mean BMI of 36.1 kg/m^2 from the overall study sample. This implies that participants with extremely high BMI values were less likely to respond than other participants.

Variables Associated With Higher BMI
The original analysis compared demographic characteristics across BMI categories after excluding those with missing data and those with BMI less than 25.0 kg/m^2. These analyses found that being younger, female, nonwhite, and not married/living as married and having some college education were all associated with a higher BMI. These results were consistent across the imputation models (Table 3).

This suggests that these associations are not greatly influenced by the missing values. The summaries of respondent characteristics from the imputed data sets are similar to those provided for those with a valid BMI presented in Table 1 (data not shown).

Logistic Model for Feeling Judged
In the original analysis, the odds of feeling judged differed significantly across BMI categories (overall $P<.001$), with those having higher BMI being more likely to report feeling judged (OR, 2.38, 4.62, and 5.26 for those with BMI 30.0-34.9, 35.0-39.9, and ≥40.0 kg/m^2, respectively). These findings remained consistent across all imputation methods (Table 4).

Logistic Model for Not Always Treated With Respect
In the original analysis, the likelihood of not always feeling respected did not differ significantly across BMI categories (overall $P=.11$), though there was some evidence that those in the highest BMI category were more likely to report this perception (OR, 1.51; $P=.03$). These findings were consistent across all imputation methods (Table 5).

Logistic Model for Not Always Treated as an Equal
Similar to the results for not always being treated with respect, in the original analysis

the likelihood of not always being treated as an equal did not differ significantly across BMI categories (overall $P=.12$) but there was some evidence suggesting that those in the highest BMI category were more likely to have this perception (OR, 1.37; $P=.02$). These findings were also relatively consistent across all imputation methods (Table 6).

Missing Not at Random Sensitivity Analyses

When looking at the effects of the missing data not being MAR, there are an infinite number of possible missing mechanisms that can be explored. Missing values can be filled in with a range of possible values to determine where the study results change. These sensitivity analyses can also assume that a proportion of the data is MAR and the rest are not MAR. Although it is impossible to explore the effects of all possible missing data options, we looked at changing the log odds in the logistic models over a wide range of options to determine whether the study results changed. Most of the study conclusions did not change in these sensitivity analyses. For the logistic model of feeling judged, participants with missing values would need to have a 6 times higher log odds of being judged than do other participants before the association with BMI would become nonsignificant. For the logistic model of being treated as an equal, the log odds would have to be 4 times higher for the association to become nonsignificant. For the respect model, the log odds would have to be 5 times higher before the association becomes nonsignificant.

DISCUSSION

The results of the sensitivity analyses of missing values support the results published in the original article. That is, relative to patients with lower BMI, participants with higher BMI were more likely to feel judged because of their weight and were less likely to feel that they were always treated with respect and as equals. There were no indications that the missing data biased these original results.

Each of the different imputation methods has their own strengths and limitations, yet they each contribute to an overall understanding of the missing data issues. Multiple imputation is widely accepted as the standard for imputation because it provides valid estimates and tests when the data are MAR or MCAR. It takes into account the variability of missing values to arrive at precise statistical tests and unbiased estimates. However, it is not a perfect solution because the results are highly dependent on applying the appropriate models to fill in the missing data. Although single imputation methods do not provide precise tests and estimates, they can be useful for sensitivity analyses. The single imputation methods make different assumptions about missing values and the resulting analyses can reveal how vulnerable the analyses are to these assumptions. Sensitivity analyses, that make reasonable assumptions about possible MNAR mechanisms, are important to determine the stability of study results.

Replacing missing values with the mean, median, minimum, and maximum make overly simplistic assumptions that missing values either are all at the extreme values or are expected to be similar to an "average" value. These methods also do not account for missing values in other variables such as age and outcome variables. The expectation-maximization algorithm, MCMC, and random forest methods avoid this problem by imputing all variables that have missing values. In particular, random forest imputation is a machine learning technique that can fill in its best guess at the missing values using all the known information, even if the unknown values are related to the known values through nonlinear relationships and interactions.

There is a strong limitation to this sensitivity analysis of missing data. The survey response rate is much lower than the typical mail survey rate of about 50%.[10] With a survey return rate of only 14%, it is likely that there are differences between individuals that returned the surveys and individuals that refused or did not return the surveys. When entire surveys are missing, this is considered unit-level missing data. The effect of this huge unit-level bias cannot be assessed using the available survey data. Some options for evaluating the scope of this unit-level nonresponse bias are (1) conducting follow-up surveys with individuals who did not respond, (2) comparing responders with nonresponders by using information available for all individuals, (3) comparing survey results with other data sources, and (4) comparing early and late responders by using

the assumption that late responders will be more similar to nonresponders.[11] From our nonresponse bias analysis, it appears that individuals with extreme BMI levels were less likely to respond. Because it is likely that individuals with extreme BMI levels are more likely to have issues with being judged, not being treated as an equal, and not being treated with respect, this could imply that the extent of these social/emotional issues may be more pervasive than reported in this survey.

Another limitation of this study and that of the original study are that they focused only on the social/emotional aspects of high BMI. They did not look at physical health aspects associated with high BMI or the interaction between physical and social aspects. High BMI can be associated with either excess adipose or being extremely fit, such as in extreme athletes. This results in a confounded analysis because the study did not distinguish between healthy and unhealthy high BMI levels.

CONCLUSION
Although the original conclusions of this study were not changed because of these additional analyses, they did provide more confidence that the conclusions are solid and provide evidence that missing data do not bias the original results. It is imperative that the effects of missing data be explored in all studies to assess the degree to which the missing data may have biased the results. Sensitivity analyses, using single and multiple imputation, can either provide evidence that missing data did not affect the study conclusions or can provide insights into the effects of the missing data.

ACKNOWLEDGMENTS
We thank Amylou Dueck, PhD for her guidance in developing the imputation macros and Veronique Roger, MD for her leadership on the Patient-Centered Outcomes Research Institute grant. This research was supported by the Patient-Centered Outcomes Research Institute, USA (PCORI) - CDRN-1501-26638 (V. Roger, PI).

SUPPLEMENTAL ONLINE MATERIAL
Supplemental material can be found online at: http://www.mcpiqojournal.org. Supplemental material attached to journal articles has not been edited, and the authors take responsibility for the accuracy of all data.

Affiliations (Continued from the first page of this article.): Rural Health, Essential Health, Duluth, MN (I.V.H.); and Center for Healthcare Delivery Innovation, Minneapolis Heart Institute, Minneapolis Heart Institute Foundation, Minneapolis, MN (S.M.B.).

Correspondence: Address to Paul J. Novotny, MS, Division of Biomedical Statistics and Informatics, Mayo Clinic, Harwick 8, 200 First St SW, Rochester, MN 55905 (novotny@mayo.edu).

ORCID
Paul J. Novotny: https://orcid.org/0000-0002-1526-9817;
Gina L. Mazza: https://orcid.org/0000-0002-5305-6193;
David Bradley: https://orcid.org/0000-0003-4006-6760;
Ivana Croghan: https://orcid.org/0000-0003-3464-3525

REFERENCES
1. Fairclough DL. *Design and Analysis of Quality of Life Studies in Clinical Trials*. Boca Raton, FL: Chapman and Hall/CRC; 2010.
2. Perkins NJ, Cole SR, Harel O, et al. Principled approaches to missing data in epidemiologic studies. *Am J Epidemiol*. 2017; 187(3):568-575.
3. Fox-Wasylyshyn SM, El-Masri MM. Handling missing data in self-report measures. *Res Nurs Health*. 2005;28(6):488-495.
4. Croghan IT, Phelan SM, Bradley DP, et al. Needs assessment for weight management: the learning health system network experience. *Mayo Clin Proc Innov Qual Outcomes*. 2018;2(4): 324-335.
5. Harrel O, Mitchell EM, Perkins NJ, et al. Multiple imputation for incomplete data in epidemiologic studies. *Am J Epidemiol*. 2018; 187(3):576-584.
6. Little RJA, Rubin DB. *Statistical Analysis With Missing Data*. 2nd ed. Hoboken, NJ: John Wiley & Sons; 2002.
7. Little RJ. Comments on: Missing data methods in longitudinal studies: a review. *TEST*. 2009;18(1):47-50.
8. Resseguier N, Giorgi R, Paoletti X. Sensitivity analysis when data are missing not-at-random. *Epidemiology*. 2011;22(2):282.
9. Sloan JA, Dueck AC, Erickson PA, Guess H, Revicki DA, Santanello NC. Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group. Analysis and interpretation of results based on patient-reported outcomes. *Value Health*. 2007;10(suppl 2):S106-S115.
10. Millar MM, Dillman DA. Improving response to web and mixed-mode surveys. *Public Opin Q*. 2011;75(2):249-269.
11. Johnson TP, Wislar JS. Response rates and nonresponse errors in surveys. *JAMA*. 2012;307(17):1805-1806.