

Table 1. Agreement Between Self-Report and Medicare Claims for Initial Treatment of Breast Cancer

Treatment	Correct, %	Kappa (95% CI)	%			
			Sensitivity	Specificity	PPV	Negative Predictive Value
Receipt of radiation therapy	97.9	0.95 (0.94–0.96)	97.6	98.6	99.3	95.3
Receipt of chemotherapy	96.8	0.90 (0.88–0.92)	90.0	98.6	94.5	97.4
Type of surgery*	92.8	0.83 (0.81–0.85)				
BCS			99.5	79.7	90.3	98.7
Mastectomy			79.7	99.5	98.7	90.3
Receipt of breast reconstructive surgery	99.1	0.90 (0.86–0.94)	90.2	99.6	90.8	99.5

Excluding missing values: radiation therapy (n = 3), chemotherapy (n = 6), type of breast surgery (n = 25), and reconstructive surgery (n = 5).

*“Best case” scenario for type of breast surgery (see text for explanation): proportion correct = 98.2%, kappa = 0.96 (95% confidence interval (CI) = 0.95–0.97), sensitivity for mastectomy = 95.9%, positive predictive value (PPV) for mastectomy = 98.9, PPV for breast-conserving surgery (BCS) = 97.9.

approximately 30 months after the surgery.⁷ Self-reported data regarding treatment (breast surgery, radiation therapy, chemotherapy, and reconstructive surgery) were validated against Medicare claims data from 1 month before diagnosis until the time of the survey.⁸ Agreement was assessed using the Cohen kappa statistic.⁹

The mean age of the women when they completed the survey was 76.1 ± 5.5 (range 67–93); 94% were Caucasian, and the majority were healthy (65% with no comorbidities). Half were married, and 92% had at least a high school degree. Median annual household income in 2004 was \$29,000. According to Medicare claims, 66% underwent breast-conserving surgery (BCS), 67% underwent radiation therapy, and 21% received chemotherapy. Only 4.6% (n = 143) underwent breast reconstruction surgery.

Assuming Medicare claims as the criterion standard, the validation of each item on the self-report questionnaire is summarized in Table 1. Overall, agreement was excellent for the four treatments examined. Kappa values varied between 0.83 for type of breast surgery and 0.95 for receipt of radiation therapy. Sensitivity (80–99%), specificity (98–99%), positive and negative predictive values (90–99%) of self-report were high for all four treatments. Proxy response, patient age group, and educational status did not significantly affect accuracy of self-report (data not shown).

Agreement for type of breast surgery was probably higher than what is reported. Of the 2,207 women who self-reported undergoing BCS, 214 had Medicare claims for mastectomy; 171 (80%) of these 214 women also had claims for BCS. Therefore, these 171 women presumably underwent initial BCS followed by mastectomy at a later date and were identified by the Medicare claims algorithm⁸ as total mastectomy cases. The survey did not capture these 171 women as undergoing eventual mastectomy, because a positive response to the BCS survey item was erroneously programmed to skip the next item about mastectomy. Therefore, these 171 women were placed in the BCS self-report group. If these 171 women are recategorized as self-reported total mastectomy cases, then the sensitivity of self-report for mastectomy increases from 79.7% to 95.9%, and the kappa improves from 0.83 to 0.96 (95% confidence interval = 0.95–0.97). These values would represent the “best case” scenario, and the true agreement probably falls somewhere between those reported in Table 1 and this “best case” scenario.

These results confirm the validity of self-report described in previous smaller studies in elderly breast cancer survivors⁶ and younger women.^{4,5} Several factors may explain the disagreement between self-report data and Medicare claims in the current study. From a survey standpoint, errors may be attributed to biased self-report, recall, or unclear wording of the survey items. It was attempted to limit this bias by providing brief descriptions of each type of treatment, referencing the time frame of treatment, and wording questions at the eighth-grade level. Although it was assumed that the Medicare claims data were correct, there are limitations of Medicare data regarding accuracy of coding, completeness of claims, and the potential for underreporting due to care provided by other agencies.^{1,3}

In summary, this study demonstrates that a population-based cohort of more 3,000 elderly breast cancer survivors accurately reported general information regarding several key cancer treatments 2 to 3 years later. Details regarding stage of disease and treatment would probably still need to be obtained from other sources, but for investigators performing survey studies in breast cancer survivors that require only broad key treatment information (quality-of-life or patient satisfaction studies), patient self-report appears to be an excellent, lower-cost alternative for obtaining this.

Tina W. F. Yen, MD, MS
Division of General Surgery

Rodney A. Sparapani, MS
Changbin Guo, BS
Division of Biostatistics

Joan M. Neuner, MD, MPH
Department of Medicine

Purushottam W. Laud, PhD
Division of Biostatistics

Ann B. Nattinger, MD, MPH
Department of Medicine
Center for Patient Care and Outcomes Research
Medical College of Wisconsin
Milwaukee, Wisconsin

ACKNOWLEDGMENTS

Conflict of Interest: The editor in chief has reviewed the conflict of interest checklist provided by the authors and has determined that the authors have no financial or any other kind of personal conflicts with this paper. This research was supported by Grants K07CA125586 (TWY), R21CA131643 and K08AG021631 (JMN), and R01CA81379 (ABN) from the National Institutes of Health, National Cancer Institute and National Institute on Aging. These grant funds are not related to products, technology, or methodologies involved in this manuscript submission.

Author Contributions: Study concept and design: Laud, Nattinger, Yen. Acquisition of subjects and data: Nattinger, Sparapani. Analysis and interpretation of data: Guo, Sparapani, Laud, Nattinger, Neuner, Yen. Preparation of manuscript: Laud, Nattinger, Neuner, Sparapani, Yen.

Sponsor's Role: None.

REFERENCES

1. Du X, Freeman JL, Warren JL et al. Accuracy and completeness of Medicare claims data for surgical treatment of breast cancer. *Med Care* 2000;38:719–727.
2. Du XL, Key CR, Dickie L et al. Information on chemotherapy and hormone therapy from tumor registry had moderate agreement with chart reviews. *J Clin Epidemiol* 2006;59:53–60.
3. Virnig BA, Warren JL, Cooper GS et al. Studying radiation therapy using SEER-Medicare-linked data. *Med Care* 2002;40:IV-49–IV-54.
4. Maunsell E, Drolet M, Ouhoumane N et al. Breast cancer survivors accurately reported key treatment and prognostic characteristics. *J Clin Epidemiol* 2005;58:364–369.
5. Phillips KA, Milne RL, Buys S et al. Agreement between self-reported breast cancer treatment and medical records in a population-based Breast Cancer Family Registry. *J Clin Oncol* 2005;23:4679–4686.
6. Schootman M, Jeffe DB, West MM et al. Self-report by elderly breast cancer patients was an acceptable alternative to Surveillance, Epidemiology, and End Results (SEER) abstract data. *J Clin Epidemiol* 2005;58:1316–1319.
7. Yen TW, Fan X, Sparapani R et al. A contemporary, population-based study of lymphedema risk factors in older women with breast cancer. *Ann Surg Oncol* 2009;16:979–988.
8. Nattinger AB, Laud PW, Bajorunaite R et al. An algorithm for the use of Medicare claims data to identify women with incident breast cancer. *Health Serv Res* 2004;39:1733–1749.
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174.

SELF-REPORTED VERSUS MEASURED HEIGHT AND WEIGHT IN THE HEALTH AND RETIREMENT STUDY

To the Editor: The global obesity epidemic has posed a major threat to population health.¹ Given that weight change and obesity have been linked with health outcomes, physical functioning, and quality of life,^{2,3} accurate population surveillance data on the trend of weight change and obesity has remained important for public health efforts to combat the obesity epidemic. In large-scale population-based studies, self-reported height and weight have been used to derive body mass index (BMI) for classifying obesity categories.^{2–5} Despite the cost savings and the general high correlation between self-reported and measured height and weight, important differences may have continued to exist in the estimated obesity prevalence, but the most recent study on the reporting accuracy of height, weight, and BMI

measures used data from 1988 to 1994 from the National Health and Nutrition Examination Survey.⁶ Therefore, the concordance between self-reported versus measured height, weight, and BMI was assessed using the 2006 data from a nationally representative sample of U.S. adults.

METHODS

The data collected from 6,799 participants of the Health and Retirement Study (HRS), a biennial longitudinal survey of U.S. adults aged 51 and older were examined.⁷ Information on self-reported height and weight was collected during interviews at each wave from 1992 to 2006. In 2006, the HRS also collected physical measures of height and weight.⁸ Height was measured without shoes and recorded in inches to the nearest quarter inch. Weight was measured using a Healthometer 830 kiloliter scale (without shoes and with light clothing; Jarden Corporation, Rye, NY) and recorded to the nearest half pound.⁸ Trained interviewers took all measurements. Mean differences between the self-reported and measured height and weight were evaluated. In addition, sensitivity and specificity of obesity classification based on self-reported data were assessed.

RESULTS

On average, respondents (men and women) overreported their height by 1% to 2% and underreported their weight by 1% to 3%, resulting in an underestimation of BMI of approximately 4%. Mean reporting errors of height tended to increase slightly with age for men and women ($P < .001$), whereas underreporting of weight decreased with age for both sexes ($P < .001$). As a result, the underestimation of BMI increased slightly with age (Table 1).

The prevalence of obesity was underestimated by 6% to 12% for all age and race and ethnicity groups and both sexes. The prevalence of overweight was overreported by 1% to 5% for all but women aged 80 and older. Obesity classification based on self-reported measures demonstrated high specificity (97–99%) across all age, sex, and racial and ethnic groups, but sensitivity was low and tended to decrease with age. Sensitivity decreased from 76% for men aged 50 to 59 to 46% for those aged 80 and older and from 83% to 57% in women.

DISCUSSION

As expected, older age was associated with decreased concordance between self-reported and measured height, whereas reporting accuracy of weight increased as age increased. The overreporting of height may have occurred because of the onset of diseases (such as osteoporosis) during the study period.⁹ In addition, the pattern of overreported height and underreported weight might indicate that social desirability affects reporting accuracy of height and weight.¹⁰ Although the average absolute reporting errors for height and weight were small, population classification based on self-reported measures consistently underestimated the prevalence of obesity. With respect to the effect of race and ethnicity on the concordance between self-reported and measured height and weight, the current analysis did not indicate a significant difference in reporting bias between racial and ethnical subgroups.