Title

Validity of self-report for fractures among a multiethnic cohort of postmenopausal women: results from the Women's Health Initiative observational study and clinical trials

Journal

Menopause The Journal of The North American Menopause Society, 11(3)

ISSN

1072-3714

Authors

Chen, Zhao
Kooperberg, Charles
Pettinger, Mary B
et al.

Publication Date

2004-05-01

DOI

10.1097/01.gme.0000094210.15096.fd

Peer reviewed

Menopause: The Journal of The North American Menopause Society
Vol. 11, No. 3, pp. 264-274
DOI: 10.1097/01.GME.0000094210.15096.FD

Validity of self-report for fractures among a multiethnic cohort of postmenopausal women: results from the Women's Health Initiative observational study and clinical trials

Zhao Chen, MPH, PhD,[1] Charles Kooperberg, PhD,[2] Mary B. Pettinger, MS,[2] Tamsen Bassford, MD,[1] Jane A. Cauley, DrPH,[3] Andrea Z. LaCroix, MPH, PhD,[2] Cora E. Lewis, MD, MSPH,[4] Simon Kipersztok, MD,[5] Carolyn Borne, RN, MN,[6] and Rebecca D. Jackson, MD,[7]

Abstract

Objective: The purpose of this study is to examine the validity of, and factors associated with, the accuracy of self-report (participant-report and proxy-report) for fractures.

Design: Study participants were from the Women's Health Initiative Clinical Trial and Observational Study cohorts. All women were postmenopausal; populations included American Indian, Asian/Pacific Islander, black, Hispanic, and non-Hispanic white. The average length of follow-up was 4.3 years. Self-reported fractures were adjudicated by reviewing medical records. The first adjudicated self-report of fractures for each participant was included in the analysis (n = 6,652).

Results: We found substantial variations in validity of self-report by the fracture site. Agreements between self-reports for single-site fractures and medical records were high for hip (78%) and forearm/wrist (81%) but relatively lower for clinical spine fractures (51%). The average confirmation rate for all single-site fractures was 71%. Misidentification of fracture sites by participants or proxy-reporters seemed to be a cause of unconfirmed self-reports. Higher confirmation rates were observed in participant-reports than in proxy-reports. Results of the multivariate analysis indicated that multiple factors, such as ethnicity, a history of osteoporosis or fractures, body mass index, years since menopause, smoking status, and number of falls in the past year were significantly ($P < 0.05$) related to the validity of self-report.

Conclusion: The validity of self-reports for fracture varies by fracture sites and many other factors. The assessed validity in this study is likely conservative because some of the unconfirmed self-reports may be due to poor medical record systems. The validity of self-reports for hip and forearm/wrist fractures is high in this study, supporting their use in epidemiological studies among postmenopausal women.

Keywords: Women's Health Initiative – Self-report for fracture – Validity – Clinical trial – Observational study – Postmenopausal women.

Received June 2, 2003; revised and accepted August 13, 2003.

From the [1]University of Arizona, Tucson, AZ; [2]Fred Hutchinson Cancer Research Center, Seattle, WA; [3]University of Pittsburgh, Pittsburgh, PA; [4]University of Alabama at Birmingham, Birmingham, AL; [5]University of Florida, Gainesville, FL; [6]UCLA Medical Center, Los Angeles, CA; and [7]Ohio State University, Columbus, OH.

The research upon which this article is based was performed pursuant to Contract No: N-01-WH-3-2115 with the National Institutes of Health, Department of Health and Human Services. Zhao Chen, MPH, PhD, is supported by a career development award (1 K01 AR02060) from the National Institutes of Health, Department of Health and Human Services.

Address correspondence to: Zhao Chen, MPH, PhD, Epidemiology and Biostatistics, College of Public Health, University of Arizona, 2501 E. Lee St., Tucson, AZ 85716. E-mail: zchen@u.arizona.edu.

Systematic review of medical records is often prohibited by regulations, technical difficulties, and high cost, especially in population-based research. Hence, ascertaining health information from participant-reports or proxy-reports is widely used in large epidemiologic studies. Validity of these reports directly affects the accuracy of measurements on exposures, outcomes, or other important covariates and ultimately determines our ability to precisely assess associations between risk factors and disease outcomes. Therefore, validation studies play an important role in epidemiological research.

Fractures are a major health problem around the world. In general, previous studies have indicated good agreement between self-reports of fractures and medical records.[1-6] However, the validity of self-reports may vary, depending on skeletal sites,[3,4,6] age, education, previous history of osteoporosis and risk of falls.[4] In the Nurses' Health Study, a perfect agreement between self-reports of fractures and medical records was reported,[1] but results from other studies suggested either overreporting[2,4-6] or underreporting.[4-8] These findings indicate that validity of self-reports for fractures is related to participants' characteristics.

Studies on the validity of self-reports for fractures are lacking among minorities and participants in clinical trials. The purpose of this study was to examine the validity of self-reported fractures and factors associated with the validity of these reports among multiethnic participants from the Women's Health Initiative (WHI) Observational Study (OS) and Clinical Trials (CT).

METHODS

Participants

As of August 31, 2001, 161,809 participants had enrolled in either the OS (N = 93,676) or CT (N = 68,133) of the WHI, and had been followed for an average length of 4.3 years (0-7.6). Detailed descriptions of the WHI study methods can be found in a previous publication.[9] Briefly, WHI is a large, nationwide, longitudinal study investigating factors that contribute to the development of major health outcomes affecting postmenopausal women, including cancer, heart disease, and osteoporotic fractures. There are four clinical trials and one observational study in the WHI. The CT arms include two hormone therapy trials (estrogen trial and estrogen plus progestin trial), a low-fat diet modification trial, and a calcium-vitamin D trial. Participants were recruited from communities where the 40 WHI clinical centers are located. Eligibility for the WHI included being postmenopausal, 50 to 79 years of age, unlikely to move or die within three years, and not currently participating in any other clinical trials. The WHI cohort consists of non-Hispanic white, Hispanic, black, American Indian, and Asian/Pacific Islander women. The WHI study was reviewed and approved by the Human Subjects Review Committee at each participating institution.

Procedure

At baseline, women completed self-administered or interviewer-administered questionnaires for eligibility screening and collection of baseline characteristics (such as demographic, reproductive, and health status information). Physical examinations were conducted, and a blood specimen was collected. The WHI OS women visited the WHI clinical center at year 3 to get additional physical measurements and to provide a blood specimen. The WHI CT women visited the clinic every year for follow-up measurements. Additional follow-up data collection was done through mail and telephone questionnaires. The CT participants completed questionnaires every 6 months and OS participants every year.

Fracture assessments

During the follow-up, 160,703 women (CT = 67,838 and OS = 92,865) had completed at least one medical history update, which included questions about fractures occurring in the previous 12 months for OS and 6 months for CT. The initial question on fracture was phrased like this: "Since (last reporting date), has a doctor told you that you had a broken, fractured, or crushed bone?" If people selected "yes," they were asked to answer a question, "Which bone did you break, fracture, or crush?" by marking all that apply from the following list: 1) hip, 2) upper leg (not hip), 3) pelvis, 4) knee (patella), 5) lower leg or ankle, 6) foot (not toe), 7) tailbone (coccyx), 8) spine or back (vertebra), 9) lower arm or wrist, 10) hand (not finger), 11) elbow, 12) upper arm or shoulder, 13) other (specify). Additional questions were asked regarding whether fractures were diagnosed or treated during an overnight hospital stay, and if an x-ray or imaging scan (MRI) was taken at the same medical facility where fractures were treated.

Fracture events were reported by either WHI participants or proxy respondents, which include family members, friends, health care professionals, and other informants. Proxy-reports were obtained when tracing participants with missed follow-up contacts or those who had died. Approximately 79% of the fracture reports were from the participants themselves. For our purposes, we defined reports from WHI participants as "participant-reports" and reports from sources other than the participants as "proxy-reports." The term "self-report" was used to refer to both proxy- and participant-reports. All self-reported hip fractures in the WHI were reviewed against medical records for confirmation by trained adjudicators. However, fractures reported at sites other than the hip were reviewed only for the CT participants (N = 67,838) and in a subset of the OS participants (n = 7,151) who were recruited at one of the three WHI bone mineral density centers (Pittsburgh, PA; Birmingham, AL; or Tucson/Phoenix, AZ). Fractures were initially adjudicated at the WHI local center. All self-reported hip fractures (either locally confirmed or reviewed and denied) and locally confirmed upper leg fractures were further reviewed at the WHI Bone Density Center at the University of California, San Francisco. In the CT, the mean length of follow-up was 4.6 years; in the OS, it was 4.1 years. In the subset of the OS from the three bone mineral density centers, the mean length of follow-up was 4.6 years.

Fracture definitions

Hip fractures in the WHI were radiographically confirmed fractures of the proximal femur, including fractures of the femoral neck, intertrochanteric region, and greater trochanter. Fractures reported at the ribs, chest/sternum, skull/face, fingers, toes, and cervical vertebrae or neck were collected but not adjudicated in the WHI. All other clinical fractures were defined as radiographically confirmed new or acute fractures at a certain bone site: upper leg (shaft of femur), pelvis, knee (patella or tibial plateau), lower leg or ankle (tibia, fibula, or talus), foot (tarsal, metatarsal, heel, or calcaneus), tailbone (sacrum or coccyx), spine or back (thoracic or lumbar spine), lower arm or wrist (radius, ulna, or carpal), hand (metacarpal), elbow (distal humerus, upper radius, or ulna), and upper arm or shoulder or collarbone (humerus, clavicle, or scapula).

Assessments of covariates

Information on age, years since menopause, race and ethnicity, education, income, smoking, alcohol intake, use of hormone therapy, history of fracture, number of falls in last 12 months, physical activity, and nutrient intakes were assessed from baseline questionnaires. Physical function and depression at baseline were measured using the 10-item Medical Outcomes Study Scale[10] and the shortened Center for Epidemiologic Studies Depression Scale (CES-D), respectively.[11] Weight was measured to the nearest 0.1 kg on a balance beam scale, with the participant dressed in indoor clothing without shoes. Height was measured to the nearest 0.1 cm using a wall-mounted stadiometer. Body mass index (BMI) was calculated as: weight (kg)/height $(m)^2$.

Statistical analysis

Descriptive analyses of baseline characteristics were conducted separately for all WHI women with follow-up medical information (n = 160,703) and in a subset of women (n = 6,652) who had adjudicated self-reports of fractures. Only the first adjudicated fracture report of each participant was included in the subsequent analyses. To assess the accuracy of reported fractures, confirmation rates were computed using fracture as the unit of measure. These rates were simply the proportion of self-reported fractures with a fracture confirmed at the exact site, or at the exact or adjacent site, out of the total number of fractures reported at that site. These rates were computed overall, and by each fracture location. Fracture reports were further categorized into participant-reports and proxy-reports, and confirmation rates were calculated for each category. Confirmed fracture at the exact site required agreement on location between self-reports and medical records. Fractures confirmed at an adjacent site were those fractures found by medical record review to be immediately contiguous to the self-reported site. For example, the pelvis and the upper leg served as adjacent fracture sites for a self-reported fracture of the hip (Appendix 1). An unrelated fracture was confirmed at neither the exact nor an adjacent site to that reported.

We use "unconfirmed self-reports" to refer both to no fracture(s) on radiological records and to self-reports that cannot be confirmed because we were unable to obtain the medical records (no records or no access to the records). Percentages for each type of unconfirmed self-reports were calculated by skeletal site. This analysis was restricted to self-reports where a fracture was the only outcome needing adjudication to limit the possibility that a fracture was accidentally identified when reviewing medical records for other health outcomes.

There were 281 multiple-site fracture reports (fractures that occurred at more than one skeletal site simultaneously in one fracture event). Considering the small number of women who had 3 or more fractures in one event (n = 20), we restricted our analyses to only single-site and two-site fractures.

Logistic regression models were used to compute odds ratio (OR) and 95% confidence interval (95% CI) for predictors of unconfirmed fractures. Here, participants were the unit of analysis so that those participants with two fractures were considered to have an unconfirmed fracture if one or both were not confirmed at their exact skeletal sites. Factors that were either associated with accuracy of self-reports in previous studies or potential predictors for validity of fracture reports in this cohort were included in the models.

To evaluate the accuracy of reported date of fracture by the participants, we compared the difference between the participant-reported hospitalization date and that determined from the confirmation process for single-site hip fractures. Because the majority of hip fractures result in immediate hospitalization, the date of admission to the hospital for hip fracture was used as the surrogate of fracture date in this analysis. Any non-hospitalized hip fractures were excluded.

All analyses were performed using SAS, version 8.2 (SAS Institute Inc., Cary, NC).

RESULTS

Baseline characteristics of the participants

Table 1 displays selected baseline characteristics of the study participants. Compared with all WHI participants, the subgroup of WHI participants in this study were older, more likely to have had previous fractures, and had an increased number of falls in the last 12 months. Most of the women in this study were from the WHI CT cohorts because only a small percentage of WHI OS women had adjudicated self-reports for fractures.

TABLE 1. *Selected baseline characteristics*

	WHI participants (N = 160,703)		Participants in this study[a] (N = 6,652)	
	n	%	n	%
Age group at screening (y)				
50-59	53,088	33.03	1,895	28.49
60-69	72,168	44.91	2,952	44.38
70-79	35,447	22.06	1,805	27.13
Ethnicity				
White	132,942	82.73	5,886	88.48
Black	14,405	8.96	382	5.74
Hispanic	6,301	3.92	175	2.63
American Indian	699	0.43	36	0.54
Asian/Pacific Islander	4,157	2.59	99	1.49
Unknown	2,199	1.37	74	1.11
Study				
CT randomized	67,838	42.21	5,701	85.70
OS enrolled	92,865	57.79	951	14.30
U.S. region				
Northeast	36,755	22.87	1,815	27.29
South	41,569	25.87	1,624	24.41
Midwest	35,289	21.96	1,279	19.23
West	47,090	29.30	1,934	29.07
Education				
< HS	8,443	5.29	317	4.80
HS/post HS	87,969	55.16	3,888	58.90
College degree or higher	63,076	39.55	2,396	36.30
History of fracture				
Never	93,548	64.10	2,752	50.43
< Age 55	35,818	24.54	1,730	31.70
≥ Age 55	16,568	11.35	975	17.87
Number of falls in last 12 mo				
None	104,032	67.48	3,632	60.20
1-2	43,695	28.34	2,022	33.52
≥ 3	6,429	4.17	379	6.28

WHI, Women's Health Initiative; CT, WHI Clinical Trials; OS, WHI Observational Study; HS, high school.
[a]A subgroup of WHI participants whose self-reports of fractures were adjudicated.

Agreements between self-reports and medical record for fractures

The accuracy of self-reports for fractures is presented for single-site fractures (Table 2) and two-site fractures (Table 3) separately. The overall confirmation rates at the exact site were 71% for single-site fractures and 57% for two-site fractures. The confirmation rate for single-site fractures at the hip was 78.2%; 80.7% at the lower arm or write; 51.4% at the spine. Considering the adjacent sites as confirmed fractures improved the confirmation rates for all the fractures, but the improvement differed across the skeletal sites, with the greatest improvement occurring in upper leg fractures. Overall, higher confirmation rates were observed at the hip, lower arm, upper arm, and lower leg site. The fracture sites with the lowest proportion of confirmed fractures were tailbone, upper leg, hand, and spine.

Among the 261 women with self-reports for two-site fractures (the total number of fractures was 522), only 31.8% of the self-reports were confirmed exactly for both sites; 26.1% were confirmed for one exact site and one adjacent site; 23.8% confirmed for one exact site and one unrelated site; 1.1% both at the adjacent sites; 1.5% for one adjacent site and one unrelated site; and for 15.7%, neither of the two were confirmed at an exact or adjacent site (data not shown).

Comparisons of validity between participant-reports and proxy-reports for fractures

The confirmation rates of participant-reports and proxy-reports for single-site and two-site fractures are presented separately in Table 4. The accuracy of participant-reports for single-site fractures was better than that for two-site fractures. Similarly, except for the el-

TABLE 2. *Accuracy of self-reports for single-site fractures*[a]

Self-report	Total	Confirmed exact site	Confirmed adjacent fracture	Confirmed unrelated fracture	No fracture confirmed	% Confirmed exact site	% Confirmed exact or adjacent site
Hip	623	487	16	0	120	78.2	80.7
Spine or back (vertebra)	510	262	3	3	242	51.4	52.0
Hand (not finger)	220	110	39	4	67	50.0	67.7
Lower arm or wrist	1,433	1,157	65	17	194	80.7	85.3
Elbow	267	187	34	0	46	70.0	82.8
Upper arm or shoulder	497	409	6	6	76	82.3	83.5
Tailbone (coccyx)	65	28	2	1	34	43.1	46.2
Pelvis	138	95	8	4	31	68.8	74.6
Upper leg (not hip)	92	26	38	7	21	28.3	69.6
Knee (patella)	284	177	17	1	89	62.3	68.3
Lower leg or ankle	1,196	931	64	14	187	77.8	83.2
Foot (not toe)	1,046	625	63	4	354	59.8	65.8
Total	6,371	4,494	355	61	1,461	70.5	76.1

[a]Self-reports include both proxy-reports and participant-reports; N = 6,371.

TABLE 3. *Accuracy of self-reports for two-site fractures*[a]

Self-report	Total	Confirmed exact site	Confirmed adjacent site	% Confirmed exact site	% Confirmed exact or adjacent site
Hip	41	30	0	73.2	73.2
Spine or back (vertebra)	30	8	0	26.7	26.7
Hand (not finger)	32	9	9	28.1	56.3
Lower arm or wrist	113	76	12	67.3	77.9
Elbow	47	27	14	57.4	87.2
Upper arm or shoulder	57	33	6	57.9	68.4
Tailbone (coccyx)	11	0	0	0.0	0.0
Pelvis	21	14	3	66.7	81.0
Upper leg (not hip)	14	3	6	21.4	64.3
Knee (patella)	41	23	8	56.1	75.6
Lower leg or ankle	65	44	10	67.7	83.1
Foot (not toe)	50	29	10	58.0	78.0
All	522	296	78	56.7	71.6

N = 261 self-reports, indicating 522 fractures.
[a]Self-reports include both proxy-reports and participant-reports. Fracture is used as the unit of analysis (eg, for a report with two fractures, hip and spine, the confirmation of each fracture is counted separately for the corresponding fracture site).

bow and upper leg, single-site proxy-reports were more accurate than two-site proxy-reports, although there were few of these fractures. Overall, the confirmation rates were higher for participant-reports compared with proxy-reports.

Reasons for unconfirmed fractures

Among the unconfirmed self-reports for single-site fractures, 54% of them showed no fractures on the radiological records and the remainder were unconfirmed due to reasons that included no radiological records or inability to access the medical records (Table 5).

Predictors for validity of self-reported fractures

Predictors for having unconfirmed (not confirmed at the exact reported fracture site) self-reports of single-site and two-site fractures are presented in Table 6. In univariate models, factors associated with a more accurate reporting of fractures were older age (60+ y), longer time since menopause (>5 y), and having a higher education and better physical function (physical function construct >90). In contrast, being black, living in the South or Midwest regions of the United States, drinking alcohol in the past, having a BMI greater than 30, having a history of osteoporosis or fracture, having sustained more than two falls in the past year, being in poor health or suffering from depression were significantly associated with poorer accuracy of self-reports. In addition, two-site fracture reports were about five times more likely to be unconfirmed compared with single-site fractures, and proxy-reports were more likely to be unconfirmed compared with participant-reports. The confirmation rate did not significantly differ by study component (CT vs OS).

TABLE 4. *Accuracy of participant-reports versus proxy-reports*

	Single-site fractures					
	Participant-reports			Proxy-reports		
	Total (n = 4,831)	% Confirmed exact site	% Confirmed exact or adjacent site	Total (n = 1,540)	% Confirmed exact site	% Confirmed exact or adjacent site
Hip	458	79.3	81.7	165	75.2	78.2
Spine or back (vertebra)	390	52.3	53.1	120	48.3	48.3
Hand (not finger)	162	53.7	72.2	58	39.7	55.2
Lower arm or wrist	1,093	82.9	87.3	340	73.8	78.8
Elbow	211	72.5	82.9	56	60.7	82.1
Upper arm or shoulder	380	83.7	84.5	117	77.8	80.3
Tailbone (coccyx)	56	41.1	44.6	9	55.6	55.6
Pelvis	100	69.0	76.0	38	68.4	71.1
Upper leg (not hip)	71	29.6	73.2	21	23.8	57.1
Knee (patella)	216	63.9	69.9	68	57.4	63.2
Lower leg or ankle	917	80.0	84.6	279	70.6	78.5
Foot (not toe)	777	61.1	66.8	269	55.8	62.8
Total	4,831	72.3	77.6	1,540	65.1	71.6
	Two-site fractures[a]					
	Participant-reports			Proxy-reports		
	Total (n = 396)	% Confirmed exact site	% Confirmed exact or adjacent site	Total (n = 126)	% Confirmed exact site	% Confirmed exact or adjacent site
Hip	29	72.4	72.4	12	75.0	75.0
Spine or black (vertebra)	22	36.4	36.4	8	0.0	0.0
Hand (not finger)	23	30.4	60.9	9	22.2	44.4
Lower arm or wrist	87	71.3	81.6	26	53.8	65.4
Upper arm or shoulder	42	69.0	78.6	15	26.7	40.0
Tailbone (coccyx)	10	0.0	0.0	1	0.0	0.0
Pelvis	18	72.2	83.3	3	33.3	66.7
Upper leg (not hip)	12	16.7	58.3	2	50.0	100.0
Knee	34	55.9	79.4	7	57.1	57.1
Lower leg or ankle	51	66.7	86.3	14	71.4	71.4
Foot (not toe)	35	65.7	85.7	15	40.0	60.0
Total	396	59.3	75.3	126	48.4	60.3

[a]Fracture is used as the unit of analysis (for a report with two fractures, such as hip and spine, the confirmation of each fracture is counted separately for the corresponding fracture category).

In the multivariate models, after adjustment for all the other covariates, being black, living in the South, being a past smoker, having a BMI greater than 30, a history of osteoporosis or fractures, more than three or more falls in the past 12 months, a two-site fracture report, and proxy-reports were related to a lower confirmation rate ($P < 0.05$). On the other hand, higher accuracy of self-reports for women who were 10 to 15 years post-menopausal were found in comparison with reports for women who were fewer than 5 years past menopause.

TABLE 5. *Reasons for unconfirmed self-reports of single-site fractures*

Unconfirmed self-reports	No fracture on the radiological record (%)	Radiological record unobtainable[a] (%)
Hip	74 (69)	33 (31)
Spine or back (vertebra)	151 (65)	81 (35)
Hand (not finger)	37 (57)	28 (43)
Lower arm or wrist	93 (49)	96 (51)
Elbow	21 (51)	22 (49)
Upper arm or shoulder	32 (44)	41 (56)
Tailbone (coccyx)	15 (45)	18 (55)
Pelvis	19 (70)	8 (30)
Upper leg (not hip)	10 (56)	8 (44)
Knee (patella)	55 (63)	32 (37)
Lower leg or ankle	92 (51)	90 (49)
Foot (not toe)	177 (52)	164 (48)

N = 1,399.

[a]No radiological records or unable to access medical records.

Validity of self-reported fracture dates

The mean difference between the reported hospitalization date and the confirmed date for single-site hip

TABLE 6. *Predictors for unconfirmed self-reported fractures*[a]

	Confirmed (n = 4,577)[b] (%)	Unconfirmed (n = 2,055)[c] %	Crude OR	Adjusted[d]	
				OR	95% CI
Age at time of fracture report (y)					
50-59	15.86	20.00	1.00	1.00	
60-69	41.60	42.38	0.81	0.96	0.77-1.20
≥ 70	42.54	37.62	0.70	0.87	0.67-1.13
Ethnicity[e]					
White	90.10	84.82	1.00	1.00	
Black	4.33	8.95	2.20	1.47	1.08-2.02
Hispanic	2.60	2.73	1.11	1.14	0.71-1.84
American Indian	0.44	0.78			
Asian/Pacific Islander	1.59	1.22			
Unknown	0.94	1.51	1.25	1.29	0.86-1.92
Study component					
CT randomized	85.82	85.45	1.00	1.00	
OS enrolled	14.18	14.55	1.03	0.94	0.77-1.15
U.S. region					
Northeast	28.40	24.72	1.00	1.00	
South	22.55	28.66	1.46	1.40	1.15-1.71
Midwest	18.92	19.85	1.21	1.23	1.00-1.52
West	30.13	26.76	1.02	1.02	0.84-1.23
Education					
< HS	4.41	5.68	1.00	1.00	
HS/post HS	58.74	59.24	0.78	1.29	0.89-1.87
College degree or higher	36.85	35.08	0.74	1.38	0.93-2.03
Income					
< $35K	46.14	47.42	1.00	1.00	
$35-< $50K	22.35	20.67	0.90	0.96	0.79-1.15
$50-< $75K	17.57	17.86	0.99	0.96	0.78-1.17
≥ $75K	13.93	14.06	0.98	1.16	0.93-1.46
Smoking					
Never smoked	51.69	49.19	1.00	1.00	
Past smoker	40.61	42.06	1.09	1.19	1.03-1.38
Current smoker	7.70	8.76	1.20	1.21	0.93-1.59
Alcohol intake					
Never drinker	11.61	11.28	1.00	1.00	
Past drinker	17.49	22.27	1.31	1.23	0.94-1.61
Current drinker	70.90	66.45	0.96	1.06	0.83-1.34
BMI					
< 25	31.78	27.42	1.00	1.00	
25.0-29.9	36.90	35.29	1.11	1.04	0.88-1.24
30.0-39.9	28.44	32.11	1.31	1.22	1.01-1.48
≥ 40.0	2.88	5.18	2.08	1.68	1.15-2.46
Years since menopause					
< 5 y	10.53	12.78	1.00	1.00	
5-<10 y	15.20	14.68	0.80	0.83	0.63-1.09
10-<15 y	18.30	17.44	0.79	0.73	0.55-0.97
≥ 15 y	55.97	55.09	0.81	0.80	0.61-1.06
History of osteoporosis					
No	90.90	87.19	1.00	1.00	
Yes	9.10	12.81	1.47	1.36	1.08-1.70
History of fracture					
Never	51.67	47.65	1.00	1.00	
< age 55	31.05	33.21	1.16	1.10	0.94-1.29
≥ age 55	17.28	19.13	1.20	1.31	1.07-1.59
Falls in last 12 mo					
None	60.52	59.77	1.00	1.00	
1-2	34.18	31.77	0.94	0.87	0.75-1.02
≥ 3	5.29	8.45	1.62	1.77	1.33-2.35

TABLE 6. *Continued*

	Confirmed (n = 4,577)[b] (%)	Unconfirmed (n = 2,055)[c] %	Crude OR	Adjusted[d]	
				OR	95% CI
Family history of fracture					
No	54.08	56.45	1.00	1.00	
Yes	45.92	43.55	0.91	0.87	0.76-1.00
Physical function score					
≤ 90 (poor function)	66.95	73.18	1.00	1.00	
> 90 (better function)	33.05	26.82	0.74	0.85	0.72-1.01
Episodes/wk moderate/strenuous activity (≥ 20 min)					
No activity	18.10	18.65	1.00	1.00	
Some activity	42.78	43.07	0.98	1.08	0.89-1.31
2-< 4 episodes/wk	16.91	16.28	0.94	1.15	0.91-1.46
4+ episodes/wk	22.22	22.00	0.96	1.20	0.95-1.50
General health (self-report)					
Excellent	15.65	12.37	1.00	1.00	
Very good	40.88	35.16	1.09	0.93	0.75-1.16
Good	34.56	38.44	1.41	1.09	0.86-1.38
Fair/poor	8.91	14.03	1.99	1.28	0.94-1.75
Depression (shortened CES-D)					
No	88.49	85.69	1.00	1.00	
Yes	11.51	14.31	1.28	1.22	0.99-1.50
Fracture type					
Single-site	98.19	91.34	1.00	1.00	
Two-site	1.81	8.66	5.14	5.10	3.59-7.26
Proxy reported fracture					
No	77.78	71.48	1.00	1.00	
Yes	22.22	28.52	1.40	1.31	1.11-1.54

OR, odds ratio; CT, Clinical Trial of Women's Health Initiative; OS, Observational Study of Women's Health Initiative; HS, high school; BMI, body mass index; CES-D, Center of Epidemiological Studies Depression Scale.
[a]Self-reported single- or two-site fractures, N = 6,632.
[b]Self-reported fractures confirmed by radiological records.
[c]No fractures confirmed on radiological records or unable to obtain medical records.
[d]Adjusted for all other variables in this table.
[e]American Indian, Asian/Pacific Islanders and unknown combined for logistic regression models.

fractures was very small and similar for both OS and CT participants. More than 91% of the participant-reported dates were within 14 days of the confirmed hospitalization date, and more than 96% of the participant-reported dates differed by less than 60 days.

DISCUSSION

In this multiethnic cohort of postmenopausal women, we found substantial variation in the validity of participant and proxy fracture reports by skeletal site. This result is consistent with findings from previous studies.[3,4,6] In our study, fractures reported at spine, hand, tailbone, and upper leg were less likely to be confirmed by medical records, whereas fractures at the hip, lower arm or wrist, upper arm, and lower leg showed the highest agreement rates with medical records. A number of unconfirmed reports were due to misidentification of fracture sites by the participants or proxy-reporters. Usually, misidentified fracture sites were located adjacent to the bone where fractures actually occurred. Misidentifications for fractures that are anatomically far away from the actual fractured bone were uncommon. When considering adjacent fracture sites in medical records as acceptable confirmation, the confirmation rate increased from 28.3% to 69.6% for a single-site fracture of the upper leg and from 50.0% to 67.7% for a single-site fracture of the hand. However, including the adjacent site did not improve the agreement for spine and upper arm fractures, suggesting that misidentified fracture sites occurred more often for certain type of fractures than others. In addition, the very low confirmation rates for hand and foot fractures may be caused by confusion with finger and toe fractures, whereas finger and toe fractures were neither adjudicated in the WHI nor included in our analyses.

In our study, the unconfirmed (at the exact site) rates of self-reports of hip fractures were 22% and 27% for single-site and two-site fractures, accordingly; and unconfirmed rates of self-reports of lower arm (or wrist) fracture were 19% and 23% for single- and two-site fractures, respectively. These findings indicate good validity of self-reports for these major osteoporotic fracture sites. In contrast, validity for self-reported spine fractures was poor. The lower confirmation rate for spine fractures may be due to ambiguity in diagnosis of fractures at this skeletal site. Future epidemiological study should selectively rely on fracture data from self-reports.

Because we did not have medical records for approximately 46% of the unconfirmed self-reports, actual overreporting or false-positive rates in our study could not be assessed. Our study may underestimate the validity of self-reported fractures because some of the unconfirmed self-reports may be true fractures that had not been recorded on medical records or the medical records were not retrievable. Hence, our results represent upper bounds for false-positive reporting in this study.

The validity of self-reports for multiple fractures has not been examined separately in previous studies. In analyzing the confirmation rate of the two-site fractures, we used fracture site as the unit. For example, if a woman reported both hip fracture and hand fracture, but only hip fracture was confirmed, then we counted the hip fracture as a confirmed report, and the hand fracture as unconfirmed in the corresponding fracture site category. In general, our data showed that the validity of self-reports for single-site fractures was superior to that for two-site fractures. Overall, self-reports for two-site fractures were about 5 times more likely to be unconfirmed (for one or both fractures) than were self-reports for single-site fractures.

In the WHI, more than 20% of fractures were reported by proxy informants, such as family members, relatives, friends, health professionals, and care providers, indicating the importance of using proxies in prospective studies with older adults. However, there is a possible tradeoff when using proxy-reports because they were less reliable than the participant-reports. Our findings were different from the Study of Osteoporotic Fractures (SOF), in which there were few proxy-reports (5%) and the proxy-reports were more accurate than the participant-reports.[4] Clearly, the impact of using proxy-reports needs further investigation.

We have found that women having a history of osteoporosis or fractures, or having more than two falls in the year before the study were more likely to have unconfirmed fracture-reports, supporting similar findings as in the SOF.[4] Furthermore, we have identified several new factors related to the accuracy of self-reports for fractures, which include years since menopause, BMI, smoking, ethnicity, and region of residence. Unconfirmed fracture-reports were more common among black women, smokers, and those who had higher BMI or lived in the south region of the United States. Women who were more than 5 years post-menopausal may be less likely to have unconfirmed self-reports for fractures. There are not ready explanations for these associations. These detected predictors for validity of self-reports may be proxies of other unmeasured variables, such as quality of medical care and socioeconomic status, or they may represent residual confounding effects of other predictors. For example, the lower confirmation rate among black participants may be related to poor medical recording and retrieving systems. Since women who are further from menopause have increased risk for bone fractures, it is possible that these women and their doctors paid more attention to the diagnosis and recording of fractures, which lead to a higher confirmation rate. Interestingly, we found that membership in the CT or OS of the WHI was not related to the validity of the fracture reports, although women in the CT completed questionnaires every 6 months and women in the OS completed questionnaires every 12 months. In contrast to the SOF,[4] we did not find any significant associations between age and the confirmation rate of fracture reports after adjustment for other covariates, regardless of whether we included or excluded years since menopause. SOF is an observational study that included non-Hispanic white women over age 65, whereas our study was composed of multiethnic postmenopausal women who were 50 or older and participated in either the OS or the CT in the WHI. It is unclear what factors contribute to the different findings in the two studies, but differences in the study populations may be one of the explanations.

For the first time, this study has assessed the validity of reported fractures among multiethnic, postmenopausal women from diverse backgrounds. Although this study had the largest minority population, compared with previous studies of this kind, the results might still be driven by the white women because 83% of our study participants were white. One of the strengths of this study is that multiple factors associated with fracture confirmation rates have been closely examined. Another strength is that we separately assessed the validity of fracture reports from participant and proxy respondents. This provides valuable information for understanding the role of proxy-reports in studies with aging populations.

Lacking false-negative data is a common problem in medical research because we cannot get records from all possible healthcare providers. We acknowledge that both overreporting and underreporting of fractures may lead to misclassification and that the inability to evaluate both types of measurement error in this study is a limitation. However, although underreporting may be a significant problem in older people if they are asked to recall lifetime fractures or fractures that occurred many years ago,[8] it is less common in older women when asking about fractures that occurred in recent years.[4,6] Examining the validity of self-reports of fractures among 292 participants in the SOF,[4] the false-negative or underreporting rate was found to be zero. In another prospective study,[6] false-negative reports for fractures were only 3% at the hip and 2% at the distal forearm among older women. These results suggest that overreporting, not underreporting, is the major cause of inaccuracy of fracture reports in prospective studies, such as the WHI.

CONCLUSION

There is moderate to excellent agreement between self-reported fracture information and medical records in the WHI. Validity of self-reports for the major osteoporotic fracture sites, such as hip and forearm, is better than for other fracture sites. Multiple factors may affect the validity of fracture reports. Self-reports of single-site fractures are more reliable than reports for two-site fractures. Proxy-reports account for a significant proportion of the self-reported fractures in this study, but the validity of proxy-reports is lower than for those reports from the study participants. The finding of overall high validity of fracture data in our study supports the utility of self-reports for fractures, particularly at selective skeletal sites. Meanwhile, this study indicates the importance of assessing variations in the validity of fracture reports and causes of over- or underreporting when studying postmenopausal women from diverse cultural and geographic backgrounds.

Acknowledgments: The WHI investigators included: Jacques E. Rossouw, Linda Pottern, Shari Ludlam, Joan McGowan, and Nancy Morris (Program Office, National Heart, Lung, and Blood Institute, Bethesda, MD); Ross Prentice, Garnet Anderson, Andrea LaCroix, Ruth E. Patterson, Anne McTiernan (Fred Hutchinson Cancer Research Center, Seattle, WA); Sally Shumaker, Pentti Rautaharju (Bowman Gray School of Medicine, Winston-Salem, NC); Evan Stein (Medical Research Labs, Highland Heights, KY); Steven Cummings (University of California at San Francisco, San Francisco, CA); John Himes (University of Minnesota, Minneapolis, MN); Bruce Psaty (University of Washington, Seattle, WA); Sylvia Wassertheil-Smoller (Albert Einstein College of Medicine, Bronx, NY); Jennifer Hays (Baylor College of Medicine, Houston, TX); JoAnn Manson (Brigham and Women's Hospital, Harvard Medical School, Boston, MA); Annlouise R. Assaf (Brown University, Providence, RI); Lawrence Phillips (Emory University, Atlanta, GA); Shirley Beresford (Fred Hutchinson Cancer Research Center, Seattle, WA); Judith Hsia (George Washington University Medical Center, Washington, DC); Rowan Chlebowski (Harbor-UCLA Research and Education Institute, Torrance, CA); Cheryl Ritenbaugh (Kaiser Permanente Center for Health Research, Portland, OR); Bette Caan (Kaiser Permanente Division of Research, Oakland, CA); Jane Morley Kotchen (Medical College of Wisconsin, Milwaukee, WI); Barbara V. Howard (MedStar Research Institute/Howard University, Washington, DC); Linda Van Horn (Northwestern University, Chicago/Evanston, IL); Henry Black (Rush-Presbyterian St. Luke's Medical Center, Chicago, IL); Marcia L. Stefanick (Stanford Center for Research in Disease Prevention, Stanford University, Stanford, CA); Dorothy Lane (State University of New York at Stony Brook, Stony Brook, NY); Rebecca Jackson (The Ohio State University, Columbus, OH); Cora Beth Lewis (University of Alabama at Birmingham, Birmingham, AL); Tamsen Bassford (University of Arizona, Tucson/Phoenix, AZ); Maurizio Trevisan (University at Buffalo, Buffalo, NY); John Robbins (University of California at Davis, Sacramento, CA); Allan Hubbell (University of California at Irvine, Orange, CA); Howard Judd (University of California at Los Angeles, Los Angeles, CA); Robert D. Langer (University of California at San Diego, LaJolla/Chula Vista, CA); Margery Gass (University of Cincinnati, Cincinnati, OH); Marian Limacher (University of Florida, Gainesville/Jacksonville, FL); David Curb (University of Hawaii, Honolulu, HI); Robert Wallace (University of Iowa, Iowa City/Davenport, IA); Judith Ockene (University of Massachusetts/Fallon Clinic, Worcester, MA); Norman Lasser (University of Medicine and Dentistry of New Jersey, Newark, NJ); Mary Jo O'Sullivan (University of Miami, Miami, FL); Karen Margolis (University of Minnesota, Minneapolis, MN); Robert Brunner (University of Nevada, Reno, NV); Gerardo Heiss (University of North Carolina, Chapel Hill, NC); Lewis Kuller (University of Pittsburgh, Pittsburgh, PA); Karen C. Johnson (University of Tennessee, Memphis, TN); Robert Schenken (University of Texas Health Science Center, San Antonio, TX); Catherine Allen (University of Wisconsin, Madison, WI); Gregory Burke (Wake Forest University School of Medicine, Winston-Salem, NC); Susan Hendrix. (Wayne State University School of Medicine/Hutzel Hospital, Detroit, MI).

REFERENCES

1. Colditz GA, Martin P, Stampfer MJ, et al. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. *Am J Epidemiol* 1986;123:894-900.
2. Bush TL, Miller SR, Golden AL, Hale WE. Self-report and medical record report agreement of selected medical conditions in the elderly. *Am J Public Health* 1989;79:1554-1556.
3. Beard CM, Melton LJ III, Cedel SL, Richelson LS, Riggs BL. Ascertainment of risk factors for osteoporosis: comparison of interview data with medical record review. *J Bone Miner Res* 1990;5: 691-699.
4. Nevitt MC, Cummings SR, Browner WS, et al. The accuracy of self-report of fractures in elderly women: evidence from a prospective study. *Am J Epidemiol* 1992;135:490-499.
5. Honkanen K, Honkanen R, Heikkinen L, Kroger H, Saarikoski S. Validity of self-reports of fractures in perimenopausal women. *Am J Epidemiol* 1999;150:511-516.
6. Ismail AA, O'Neill TW, Cockerill W, et al. Validity of self-report of fractures: results from a prospective study in men and women across

Europe. EPOS Study Group. European Prospective Osteoporosis Study Group. *Osteoporos Int* 2000;11:248-254.
7. Akesson K, Gardsell P, Sernbo I, Johnell O, Obrant KJ. Earlier wrist fracture: a confounding factor in distal forearm bone screening. *Osteoporos Int* 1992;2:201-204.
8. Jonsson B, Gardsell P, Johnell O, Redlund-Johnell I, Sernbo I. Remembering fractures: fracture registration and proband recall in southern Sweden. *J Epidemiol Community Health* 1994;48:489-490.
9. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin Trials* 1998;19:61-109.
10. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-483.
11. Burnam MA, Wells KB, Leake B, Landsverk J. Development of a brief screening instrument for detecting depressive disorders. *Med Care* 1988;26:775-789.

APPENDIX

Appendix 1. *Definition of adjacent fractures*

Primary fracture site	Adjacent fracture site(s)	
Hand (not finger)	Lower arm	
Lower arm or wrist	Hand	Elbow
Elbow	Lower arm	Upper arm
Upper arm or shoulder	Elbow	
Spine or back (vertebra)	Tailbone	
Tailbone (coccyx)	Spine	Pelvis
Pelvis	Tailbone	Hip
Hip	Pelvis	Upper leg
Upper leg (not hip)	Hip	Knee
Knee (patella)	Upper leg	Lower leg
Lower leg or ankle	Knee	Foot
Foot (not toe)	Lower leg	