

# Validation of Self-Report of Rheumatoid Arthritis and Systemic Lupus Erythematosus: The Women's Health Initiative

BRIAN T. WALITT, FLORINA CONSTANTINESCU, JAMES D. KATZ, ARTHUR WEINSTEIN, HONG WANG, ROHINI K. HERNANDEZ, JUDITH HSIA, and BARBARA V. HOWARD

**ABSTRACT. Objective.** The Women's Health Initiative (WHI), initiated in 1993, enrolled 161,808 postmenopausal women aged 50–79 years and followed them with annual questionnaires for 8 years in order to study major causes of morbidity and mortality. Our objective was to determine the most effective and efficient means to validate self-reported rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) in the WHI.

**Methods.** Data from 2 of 40 WHI clinical centers were used. Of these 7443 women, 643 self-reported RA and 106 self-reported SLE. Research coordinators contacted these women using mailers and telephone calls to obtain medical record releases and a Connective Tissue Screening Questionnaire (CSQ). Medical records were obtained on 286 self-reported RA and 34 self-reported SLE and reviewed by 3 rheumatologists blind to the self-reported diagnoses. Sensitivity, specificity, and the kappa statistic were computed to evaluate the level of agreement between self-report and chart review.

**Results.** Self-reported RA was accurate only 14.7% (42/286 cases) of the time. Coupling the self-report to medication data improved the positive predictive value (PPV; 62.2%) and kappa (0.53), suggesting a moderate agreement to chart review. Self-reported SLE was accurate only 11.8% (4/34 cases) of the time. Coupling the self-report to medication data improved the PPV (40.0%) and kappa (0.44), suggesting a moderate agreement to chart review. The CSQ was inferior to using medication data but was substantially better than self-report alone.

**Conclusion.** The performance of disease self-report coupled with medication history in validating RA and SLE was very good and should obviate the need for time-consuming medical record reviews. (First Release April 1 2008; J Rheumatol 2008;35:811–8)

## Key Indexing Terms:

DIAGNOSTIC VALIDATION      SELF-REPORT      RHEUMATOID ARTHRITIS  
SYSTEMIC LUPUS ERYTHEMATOSUS      WOMEN'S HEALTH INITIATIVE

The causes of rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) are still not known despite impressive strides in the treatment of these diseases. Issues of disease causation and personal risk are often addressed with epidemiologic studies. One such study, the Women's Health

Initiative (WHI), enrolled 161,808 postmenopausal women at 40 centers in the United States<sup>1</sup>. While the WHI was not designed to directly study RA and SLE, it did collect information on self-reported rheumatic disease diagnoses, demographic information, an exhaustive array of environmental and dietary exposures, activity indices, medication data, serial measures of general well-being and musculoskeletal symptoms, and health morbidity and mortality information on all participating women. The WHI dataset has the potential to be a resource for investigators to examine many questions concerning the predictors, determinants, and natural history of RA and SLE.

To use the WHI dataset to study epidemiological questions about RA and SLE, it is necessary to understand how accurate the self-reported disease designations are. If the agreement between self-reported disease and actual disease is poor, it could create an ascertainment bias that could affect WHI analyses in unpredictable ways. Historically, the validity of self-reported diagnoses in RA and SLE has varied widely according to differences in geography and validation methodology (Table 1). The validity of self-reported

From the Washington Hospital Center and George Washington University Hospital, Washington, DC; and MedStar Research Institute, Hyattsville, Maryland, USA.

The WHI program is funded by the National Heart, Lung and Blood Institute, US Department of Health and Human Services. Funding for our study was provided by an American College of Rheumatology Clinical Investigator Fellowship Award and a Washington Hospital Center Graduate Medical Education research grant.

B.T. Walitt, MD, MPH, Assistant Professor of Medicine; A. Weinstein, MD, Professor of Medicine; B.V. Howard, PhD, Professor of Medicine, Georgetown University, Washington, DC; J.D. Katz, MD, Associate Professor of Medicine; J. Hsia, MD, Professor of Medicine, George Washington University; F. Constantinescu, MD, Assistant Professor of Medicine, Virginia Commonwealth University, Richmond, Virginia; H. Wang, MD; R.K. Hernandez, MPH, MedStar Research Institute.

Address reprint requests to Dr. B. Walitt, Washington Hospital Center, 110 Irving Street, Washington, DC 20010. E-mail: Brian.t.walitt@medstar.net  
Accepted for publication January 9, 2008.

Table 1. Historical confirmation of self-reporting of rheumatic disease.

Article	Location	Population	Self-report RA: PPV, %	Self-report SLE: PPV %	Confirmation
Rasooly <sup>2</sup>	Manchester, UK	University rheumatology outpatient	90% sensitivity*	—	Self-report compared to clinic diagnostic registry
Johnson <sup>3</sup>	Birmingham, UK	Outpatient and support groups	—	34	Suspected SLE cases from the community compared to chart review
Hochberg <sup>4</sup>	USA	General	—	33	Self-report compared to chart review
Kvien <sup>5</sup>	Oslo, Norway	General	22	—	Self-report compared to rheumatologic evaluation
Star <sup>5</sup>	Baltimore, USA	Study of osteoporotic fractures	21	—	Self-report compared to chart review and radiographic review
Barlow <sup>7</sup>	Manchester, UK	Outpatient and general	96	—	Postal surveys compared to general practitioner chart review
Karlson <sup>8</sup>	USA	Women's health cohort	35.8	20.8	Self-report compared to chart review
Ling <sup>9</sup>	Baltimore, USA	Women's Health and Aging Study	34	—	Self-report of seropositive women compared to chart review, medication review, physician evaluation, and hand photograph algorithm
McAlindon <sup>10</sup>	USA	Black Woman's Health Study	—	59–84	Self-report ± medications compared to chart review
Cerhan <sup>11</sup>	USA	Iowa Woman's Health Study	13.3	—	Self-report compared to chart review
Costenbader <sup>12</sup>	USA	Nurses' Health Study	7	—	Self-report compared to chart review and CSQ

\* Positive predictive value not reported. PPV: positive predictive value; RA: rheumatoid arthritis; SLE: systemic lupus erythematosus; CSQ: Connective Tissue Screening Questionnaire.

diagnoses has been reported to be between 7% and 96% in RA and 21% and 84% in SLE, with validations from cohorts of independent-living women similar to the WHI tending to have lower confirmation rates<sup>2–12</sup>.

Prior studies make it clear that the validity of self-reported diagnosis varies widely from cohort to cohort. Therefore, our primary goal was to validate self-reported RA and SLE diagnostic information collected by the WHI. As prior studies suggest that strict reliance on self-reported diagnoses can be problematic, we also wished to determine if the use of other data collected by the WHI, such as medication data and/or a disease screening questionnaire, could improve diagnostic accuracy. Our study design uses data collected from 2 WHI clinical centers and compares incrementally sophisticated validation methods to the results of medical chart reviews and physician interviews to determine the most effective means to define RA and SLE.

## MATERIALS AND METHODS

The WHI clinical trials and observational study, initiated in 1993, were designed to study the major causes of morbidity and mortality in postmenopausal women. The WHI observational study enrolled 93,676 postmenopausal women aged 50–79 years from 40 clinical centers across the US and tracked their health for an average of 8 years. Its purpose was to give reliable estimates of the extent that known risk factors predict heart disease, cancers, and fractures and to identify new risk factors for these health problems. The WHI clinical trials included trials of the effects of postmenopausal hormone therapy on coronary heart disease, hip fracture, and breast cancer, using separate trials of estrogen plus progestin and estrogen alone<sup>1</sup>. The trials randomized 16,608 women with an intact uterus to

0.625 mg of conjugated equine estrogens plus 2.5 mg of medroxyprogesterone or placebo and prospectively monitored the women for an average of 5.6 years; and randomized 10,739 women without an intact uterus to 0.625 mg of conjugated equine estrogens or placebo and prospectively monitored the women for an average of 7.1 years. The dietary modification WHI clinical trial randomized 48,835 postmenopausal women to either a low-fat, high fruit, vegetable and grain diet or usual eating habits and measured the effects of diet on breast cancer, colorectal cancer, and heart disease in postmenopausal women over a period of 8 to 12 years<sup>1</sup>. Further details of the WHI study design, recruitment, screening, randomization, and eligibility criteria are described elsewhere<sup>1,13–16</sup>.

Our study was a blinded validation of self-reported rheumatic diagnoses at 2 of the 40 clinical sites in the WHI. The 2 clinical sites were the MedStar Research Institute (n = 3682) and George Washington University Hospital (n = 3761), both located in the Washington, DC, metropolitan area. All women in the observational study and clinical trials who self-reported RA, SLE, or osteoarthritis (OA) were eligible to participate. Eligible women were identified from the WHI database and contacted by mail and telephone. Participants provided informed consent, physician information, and medical record releases. Medical records to support the diagnoses of RA and/or SLE were obtained and reviewed by a group of rheumatologists blinded to the self-reported diagnoses of the participants.

**Participant identification.** As part of the WHI observational and clinical trials, all participants completed a yearly health questionnaire that included the following 3 questions about arthritis: “Did your doctor ever say that you had arthritis?” (Yes/No), “What type of arthritis do you have?” (RA, Other/Osteoarthritis), and “Has a doctor told you that you have systemic lupus erythematosus (“lupus” or SLE)?” (Yes/No). Those who self-reported RA or SLE at baseline (prevalent cases) and during later followup (incident cases) were both included. A total of 643 women with self-reported RA and 106 with self-reported SLE were identified at the 2 sites. In addition, 76 women with self-reported other arthritis/OA were identified at random to serve as a control population in the review process.

**Participant contact.** Eligible women were mailed a packet of study materials containing a letter of introduction, an Institutional Review Board approved informed consent and Health Insurance Portability and Accountability Act waiver, the Connective Tissue Screening Questionnaire (CSQ)<sup>17</sup>, an Arthritis Update questionnaire, a medical record release, and a stamped envelope in which to return forms. Eligible women not responding to our initial mailing were contacted by telephone to discuss the study and their potential involvement.

**Study materials. Demographic data.** Demographic information was extracted from the WHI database. In general, the participants from our 2 sites were an average 62.4 years of age, either Caucasian (54.7%) or African American (38.6%), and well educated (91% completed high school). The WHI collected demographic information only at the initial screening visit.

**Medications.** The WHI database had collected a medication inventory at baseline and at study year 3. Additionally, all participants in the controlled trials had medication information collected at study years 1, 6, and 9. To do so, women were instructed to bring all medications they were currently taking, including over-the-counter and herbal medications, to visits to be recorded by research staff. Medications used at any time during the study, including at baseline, were included in our analyses. Only medication names and pill strength were collected by the WHI staff.

**Medication definitions.** The following medications were defined as medications used in the treatment of RA: hydroxychloroquine, sulfasalazine, minocycline, methotrexate, leflunomide, azathioprine, cyclosporine, gold, cyclophosphamide, antirheumatic biologic agents (i.e., tumor necrosis factor- $\alpha$  and interleukin 1 antagonists), and oral steroids.

The following medications were defined as medications used in the treatment of SLE: hydroxychloroquine, sulfasalazine, methotrexate, leflunomide, azathioprine, cyclosporine, mycophenolate mofetil, cyclophosphamide, and oral steroids.

Because of the use of oral steroids in many other conditions, the analyses were also performed excluding oral steroids from the medication definitions. No participant reported D-penicillamine or tacrolimus use during the study.

**Connective Tissue Screening Questionnaire.** The CSQ is based on the American College of Rheumatology (ACR) classification criteria for SLE and RA and several other connective tissue diseases. The questionnaire consists of 30 items with yes/no responses. Scoring is performed using a scoring algorithm. Scores are expressed as "probable," "possible," or "no" disease categories. The CSQ has been reported to be 85% sensitive and 92% specific for detecting RA and 96% sensitive and 85% specific for detecting SLE<sup>17</sup>. The CSQ fits entirely on one page and can be optically scanned and automatically scored.

**Arthritis Update Questionnaire.** Participants were asked for the names and contact information for their physicians. The questionnaire specifically inquired about the doctor(s) that treat their arthritis. It also collected information about their primary physicians that provided them care from 1988 (5 yrs prior to subject involvement with the WHI) through the present.

**Medical chart review.** An experienced rheumatologist (BW, FC, and JK) blinded to the self-reported diagnosis of the participant contacted the appropriate physician's office to obtain a copy of the participant's medical record, obtain a physician validation questionnaire, or to interview the physician over the telephone. All these options were allowed in order to maximize community physician response rates and to fill in potential gaps in the medical record. The physician validation questionnaire queried the diagnoses of RA, SLE, and OA using the defined ACR classification criteria as a guideline for each disease<sup>18-22</sup>. Reviewers used standard questions during the interviews that followed the ACR criteria. Reviewers made requests from one physician at a time, starting with the doctor currently caring for the participant's condition. If an office did not respond to a request or the medical record was incomplete, the physician was contacted by telephone by the reviewer. After completing a review of all the available records from one community physician, the reviewer determined if, in their judgment, the outpatient chart they reviewed contained adequate informa-

tion to finish their review. If the information obtained was inadequate, the next doctor on the list was contacted and the review process repeated until the reviewer felt that adequate information was obtained or all materials were exhausted. Only then did the reviewer determine if the participant met any of the study definitions for arthritis. Each reviewer used his or her clinical impression of the case to assign the presence or absence of RA, SLE, and OA. A strict ACR criteria cutoff was not employed. If one of these rheumatic conditions was thought to be present, each reviewer used a 5-level ordinal scale to assign a level of confidence. The type of physician making the diagnosis (i.e., rheumatologist, orthopedist, general internist/family physician) was also recorded by the reviewers.

To measure potential inter-reviewer variability, 10% of the charts were randomly selected to undergo a blinded double-review process. Diagnostic differences between the reviewers were arbitrated by a separate blinded rheumatologist (AW). If diagnostic agreement between the reviewers was in excess of 95% for the RA and SLE diagnoses, it was decided that the remainder of the reviews would be performed using a single reviewer method. Otherwise, the double-reviewer method would be employed throughout the study.

**Data entry and management.** Study data were entered into a specially designed computer program by each individual reviewer. The program was designed with deliberate redundancy to minimize significant data entry errors. Each reviewer had an individual database at a local computer to store the entered data. Each month the reviewers electronically sent a copy of their database to the study statisticians. Locally collected data were then combined with previously collected WHI data requested from the central WHI database.

**Statistical analysis.** The positive predictive values (PPV) and negative predictive value (NPV) were determined for self-report alone, self-report coupled with antirheumatic use, self-report with positive score on CSQ, and self-report with antirheumatic use and positive score on CSQ, using the chart validation data as the gold standard. The kappa statistic with exact binomial 95% confidence intervals was used to evaluate the level of agreement between self-reported diagnoses and chart reviewed diagnoses. Typical statistical methods were used to examine disagreements. Tests for equal kappa coefficients were performed to determine the effect of education, age, and income on the agreement between self-reported diagnosis and chart review findings. Differences in demographics and confirmation rates between the George Washington University and MedStar Research Institute cohorts as well as between prevalent and incident cases were determined using chi-square tests and general linear models. Our sampling frame did not attempt to identify false-negative cases under the assumption that false-negative cases would be exceedingly rare. Sensitivity and specificity were calculated and are reported, but our methods make their interpretation more difficult.

## RESULTS

A total of 348 prevalent and 295 incident cases of RA and 26 prevalent and 80 incident cases of SLE were reported in 7.7 years of followup of 7443 women in the WHI at our 2 centers. Of the 825 eligible women (including OA controls), adequate records and physician interviews were obtained for 367 women, a 44% response rate (Table 2). Of the 458 women without records, 140 declined to participate, 176 did not respond, 40 could not be contacted, and 11 provided CSQ data but had incomplete or absent medical records. Twenty-five eligible women had died prior to study initiation. Completed CSQ were obtained on 272 women. Only 261 women completed both the CSQ and medical record review. The demographic characteristics of the study participants are shown in Table 3. The study cohort mainly con-

Table 2. Profile and attrition of self-reported rheumatic disease in study cohort.

Self-Reported Diagnosis	Potential Participants	Participant Chart Reviews Obtained	Total Confirmed Cases in Study by Chart Validation	Agreement Between Self-Report and Validation, n (%)
Total	764	348		
RA	643	286	42	42 (14.7)
SLE	106	34	4	2 (11.8)
OA*	684	311	241	229 (73.6)
RA only	580	266	42	42 (15.8)
SLE only	43	14	4	1 (7.1)
RA and SLE	63	20	0	0
OA controls	76	47	241	34 (72.3)

\* Includes OA controls and participants reporting both RA and OA or SLE and OA. RA: rheumatoid arthritis; SLE: systemic lupus erythematosus; OA: osteoarthritis.

sisted of elderly, well educated, and non-impovertised women. There were no demographic differences between the responder and the nonresponder groups, with the exception of education. Responders tended to be more educated than the nonresponders. Of the responders, there were significant demographic differences between the 2 sites. The George Washington University cohort was predominately Caucasian, while the MedStar site was predominately African American ( $p < 0.0001$ ). The George Washington cohort was more educated ( $p < 0.0001$ ) and had higher incomes ( $p < 0.0001$ ) than the MedStar cohort.

Thirty-two charts (~ 10%) were randomly selected by the study statistician to undergo double-review. There were no disagreements between the rheumatologists in regard to the

diagnosis of RA and SLE. Due to the high degree of agreement between reviewers in determining the diagnoses of primary interest (RA and SLE), a single-reviewer format was utilized.

The performance of the different epidemiologic definitions of RA compared to the chart review findings is shown in Table 4. Of the 286 women with self-reported RA, only 42 were confirmed by chart review (14.7%). Coupling the self-reported diagnosis to medication data improved the PPV (62.2%), NPV (93.9%), and kappa (0.53), suggesting a moderate agreement to chart review findings with few false-negative cases. The addition of the CSQ to the definition improved the PPV (82.4%) but also decreased the NPV (91.4%) and kappa value (0.49). Using either medication or

Table 3. Study cohort demographics.

Variable	All Participants, n = 7443	MRI Site Participants, n = 3682	GW Site Participants, n = 3761	Self-Reported RA or SLE, n = 749	Self-Reported RA or SLE with Chart Reviews, n = 348	Confirmed RA or SLE, n = 46
Age, mean (SD) yrs	61.7 (7.4)	61.7 (7.0)	61.6 (7.7)	62.4 (7.3)	62.6 (7.1)	61.5 (7.7)
Race/ethnicity, %						
White	64.2	41.4	86.5	54.7	56.0	69.6
Black	30.3	53.2	7.8	38.6	39.4	30.4
Hispanic	1.9	1.7	2.1	2.4	2.3	0
Education, %						
< HS diploma	4.2	7.3	1.1	7.0	6.7	0
HS diploma	15.7	21.7	9.9	20.0	11.0	4.4
Post-HS	22.5	24.5	20.6	24.6	6.4	31.1
College degree or higher	56.0	44.3	67.5	46.7	76.0	64.5
Income, %						
< \$20,000	9.6	15.0	4.3	13.9	14.9	9.3
\$20– < \$35	14.0	17.8	10.3	15.3	17.0	14.0
\$35 – < \$50	15.9	17.9	14.0	14.9	15.8	16.3
\$50 – < \$75	22.1	21.2	23.1	22.1	24.0	25.6
≥ \$75	32.0	20.3	43.5	26.3	26.1	34.8
Don't know	1.6	2.3	1.0	2.8	2.1	0

GW: George Washington University; MRI: MedStar Research Institute; RA: rheumatoid arthritis; SLE: systemic lupus erythematosus; HS: high school.



Table 4. Performance of rheumatoid arthritis definitions compared to chart review.

	Chart Review			Positive Predictive Value, %	Negative Predictive Value, %	Sensitivity, %	Specificity, %	Simple Kappa Coefficient (95% CI)
	Yes	No	Total					
Self-report								
Yes	42	244	286	14.7	100	100	20.3	0.06 (0.04–0.08)
No	0	62	62					
Total	42	306	348					
Self-report + medication*								
Yes	23	14	37	62.2	93.9	54.8	95.4	0.53 (0.39–0.67)
No	19	292	311					
Total	42	306	348					
Self-report with CSQ**								
Yes	25	39	64	39.1	94.9	71.4	82.7	0.44 (0.27–0.53)
No	10	187	197					
Total	35	226	261					
Self-report + medication + CSQ**								
Yes	14	3	17	82.4	91.4	40.0	98.7	0.49 (0.32–0.66)
No	21	223	244					
Total	35	226	261					

\* Medications include hydroxychloroquine, sulfasalazine, minocycline, methotrexate, leflunomide, azathioprine, cyclosporine, gold, cyclophosphamide, antirheumatic biologic agents (i.e., anti-tumor necrosis factor- $\alpha$  and anti-interleukin 1), and oral steroids. \*\* CSQ scored as “probable” disease; N decreases due to number of participants completing the CSQ. CSQ: Connective Tissue Screening Questionnaire.

CSQ data in conjunction with self-report did not improve the kappa value (0.42). Use of more or less stringent CSQ disease definitions (i.e., probable disease vs probable + possible disease) did not lead to significant alterations in the test’s performance. The analyses were repeated after removing oral steroids from the list of defining medications. The RA and medication definition had a greater PPV (67.9%) with a decrease in kappa (0.49). Similar findings were noted when CSQ results were also included. Overall, the addition of oral steroids to the medication definition improved the performance of RA epidemiologic definitions.

Further analyses were performed on the false-positive and false-negative participants when the self-report plus medication definition was employed. Fourteen women who self-reported RA and were taking appropriate medications who had negative chart reviews (false-positives) had the following conditions: OA (12), Sjögren’s syndrome (2), polymyalgia rheumatica (2), undifferentiated connective tissue disease (1), unspecified seronegative spondyloarthritis (1), eosinophilic fasciitis (1), and autoimmune hepatitis (1). Review of the 19 women found to be false-negative with the self-reported RA with medication definition revealed that all had self-reported RA, and that 17 had documentation of medication use in their medical records that was not identified in the WHI database. Ninety-three percent of the cases of RA confirmed by medical records had documentation from a rheumatologist. There was no difference in rates of confirmation between prevalent and incident cases ( $p = 0.13$ ) or between the 2 study sites ( $p = 0.23$ ).

The performance of the different epidemiologic definitions of SLE compared to the chart review findings is shown

in Table 5. Of the 34 women with self-reported SLE, only 4 were confirmed by chart review (11.8%). Combining the self-reported diagnosis with medication data improved the PPV (50.0%) and kappa (0.49), suggesting a moderate agreement to chart review findings with few false-negative cases. The addition of the CSQ to the definition decreased the PPV (33.3%) and kappa (0.39) without much change in NPV (99.6%). Using either medication or CSQ data in conjunction with self-report did not improve the kappa value (0.24). Use of alternative CSQ disease definitions did not lead to significant alterations in test performance. The analyses were repeated after removing oral steroids from the list of defining medications. The addition or removal of oral steroids from the medication definition had no effect on the performance of SLE epidemiologic definitions.

Further analyses were performed on the false-positive and false-negative participants when the self-report plus medication definition was employed. Three women had self-reported SLE, were taking appropriate medications, and had negative chart review for SLE. They all had OA; one subject also had positive antinuclear antibodies without symptoms and another had Sjögren’s syndrome. Review of the 2 women with positive chart reviews for SLE that did not meet the self-report criteria for SLE revealed that neither had self-reported SLE and both had documentation of medication use in their charts. One hundred percent of the confirmed cases of SLE had records from a rheumatologist. There was no difference in rates of confirmation between prevalent and incident cases ( $p = 1.0$ ) or between the 2 study sites ( $p = 1.0$ ).

Education did affect the self-reported accuracy of RA,

Table 5. Performance of systemic lupus erythematosus definitions compared to chart review.

	Chart Review			Positive Predictive Value, %	Negative Predictive Value, %	Sensitivity, %	Specificity, %	Simple Kappa Coefficient (95% CI)
	Yes	No	Total					
Self-report								
Yes	2	32	34	5.9	99.4	50.0	90.7	0.09 (−0.04–0.21)
No	2	312	314					
Total	4	344	348					
Self-report + medication*								
Yes	2	2	4	50.0	99.4	50.0	99.4	0.49 (0.07–0.92)
No	2	342	344					
Total	4	344	348					
Self-report with CSQ**								
Yes	1	5	6	16.7	99.6	50.0	98.1	0.24 (−0.16–0.64)
No	1	254	255					
Total	2	259	261					
Self-report + medication + CSQ**								
Yes	1	2	3	33.3	99.6	50.0	99.2	0.39 (−0.15–0.94)
No	1	257	258					
Total	2	259	261					

\* Medications include hydroxychloroquine, sulfasalazine, methotrexate, leflunomide, azathioprine, cyclosporine, cellcept, cyclophosphamide, and oral steroids. \*\* CSQ scored as “probable” disease; N decreases due to number of participants completing the CSQ. CSQ: Connective Tissue Screening Questionnaire.

with college educated participants more likely to self-report correctly compared to participants with less education (chi-square 10.44,  $p = 0.001$ ). This was not seen in SLE or OA. Income and age had no significant effect on self-reported accuracy for all 3 conditions, although persons with incomes lower than \$35,000 were less likely to accurately report disease (chi-square 3.52,  $p = 0.06$ ).

## DISCUSSION

Prior studies<sup>3-6,8,9,11,12</sup> have reported poor concordance between self-reported diagnoses and true cases of RA and SLE, observations that are reinforced by the ~ 15% concordance of self-reported RA cases and the ~ 12% concordance of self-reported SLE cases in the WHI. This demonstrates that there is a problematic gulf between patients and their physicians in understanding RA. While some of this misunderstanding is a function of education, it alone does not account for such poor concordance. However, all the participants with confirmed RA did report that they have the disease. No case of RA confirmed by chart review was improperly coded as SLE or OA by the patients. This observation with self-reported RA has been reported in most other validation studies as well<sup>2-6</sup>. False-negative reporting of RA does not seem to be a problem, which is an advantage in epidemiological studies. Our results provide guidance for developing valid and reproducible means to confirm self-reports of RA and SLE in an epidemiologic setting. We report that it is possible to use self-reported data available in a typical epidemiologic study to provide a reasonably specific diagnosis of RA and SLE.

The use of medication history to confirm the diagnosis of RA has not been frequent in prior validation studies. The

majority of studies describe utilizing serology, specialized RA questionnaires, and research algorithms<sup>23,24</sup>. Only one previous study utilized medication data for diagnostic confirmation, but it was used as part of a larger algorithm that included thorough examinations<sup>9</sup>. The use of medication data to validate self-reported RA diagnosis appears to be a reasonable strategy from our data. This strategy may miss about 45% of true RA cases, but 62% of the cases recognized are valid cases. Our data with this strategy enables WHI investigators to be more confident that they are studying RA. Indeed, we suspect that the use of medication data would have performed even better with a more directed inquiry about medication use during the initial data gathering. Seventeen of 19 patients who did have RA but were recorded not to be taking medication actually were using disease modifying antirheumatic drugs (DMARD) on chart review. Had these medications been identified by the WHI, the self-report and medication definition would have had a PPV of 74%. The WHI data provided a snapshot of medication use once every 3 years that focused only on medications taken at least twice a week. It is possible that this precluded some participants from reporting weekly methotrexate or periodic biologics use. A more regular review of medication use or asking the participants directly about DMARD use may have improved performance. Regardless, a diagnostic strategy of combining self-report and medication data to define cases of RA in epidemiological studies appears to be a reasonable approach.

Our data were similar in regard to validation of self-reported SLE. Self-report coupled with medication yielded the highest kappa value. However, these results should be considered carefully due to the low numbers of actual cases

and the unexpected finding of 2 cases of SLE that were not self-reported by the participants.

The use of self-report questionnaires to improve diagnostic validity in epidemiologic studies has been addressed in several studies. Our data suggest that use of a screening questionnaire for multiple autoimmune disorders does not perform much better than using medication data for validation. Our study demonstrates that questionnaire use decreased PPV and did not improve agreement with chart review. This small decrease in PPV is more likely a reflection of a decrease in sample size than an implicit characteristic of the questionnaire. The combination of both medication data and the questionnaire also did not provide any improvement in agreement with chart review. The combination of the 3 datapoints was too stringent, excluding a large number of participants with RA. Even an "either/or" paradigm for medication data and the screening questionnaire did not improve performance. These results may reflect our use of the CSQ in a role it was not designed for. The CSQ may not appear to be well suited for the aging WHI study population. Our elderly population may have had difficulty with the language and structure with the CSQ, as evidenced by 87 participants filling out all the study paperwork properly but refusing to complete the CSQ. Many others needed to be contacted directly to clarify or complete the questionnaire as well. Perhaps a questionnaire with larger print and a simpler format may have performed better. While the CSQ did not outperform prospectively collected medication data, it clearly provides a substantial benefit over self-report alone, demonstrating very high specificities.

There are some drawbacks in using our approach to defining disease in an epidemiologic study. Requiring medication use to be defined as a case is, in essence, a tautology. It assumes that all cases of the disease have been properly diagnosed and treated. This excludes undiagnosed, untreated, misdiagnosed, and mistreated cases. This would be problematic in studying early manifestations of these diseases. Using medications to define disease also requires a meticulous medication data collection. Our study suggests that nearly 40% of actual cases are missed due to inaccurate medication data collection. Both these drawbacks can be partially controlled for in studies by the use of blinding and randomization, such as the WHI controlled trials, but these drawbacks should be considered in interpreting epidemiologic studies utilizing these definitions. The investigators also chose a validation methodology that utilized reviewer impression rather than ACR criteria for RA and SLE for determining the validity of individual cases. The ACR criteria have not been validated for use in chart reviews but do represent a standard diagnostic tool that can be reproduced by other research groups. A final drawback with this study was its 44% response rate. This affects the ease with which these results can be generalized to larger patient populations.

These results highlight the difficulties of managing ascer-

tainment bias in large epidemiologic studies of rheumatic diseases. Our data suggest that despite being well educated and of fair financial means, women with various forms of rheumatic disease tend to incorrectly identify their type of arthritis. Our study demonstrates that medication histories appear to be very good for confirming self-reported cases of RA and SLE. This can obviate the need for time-consuming medical chart review. Given the large size of the WHI cohort, this concordance is likely adequate to allow many valuable analyses of the natural history and predictors of RA and SLE in postmenopausal women. We hope that our experiences in defining RA and SLE in the WHI will aid investigators working with large research datasets in performing their own diagnostic validations, and provide insight for community practitioners into how patients with rheumatic diseases truly understand their conditions.

## ACKNOWLEDGMENT

We acknowledge Karen Barr, George Washington University site coordinator; Amy Smith, MedStar Research Institute site coordinator; Kay Mickel, MedStar Research Institute site clinical manager; Donna Embersit, George Washington University site clinical manager; Mary Pettinger, Women's Health Initiative Statistical Support; and Greg Foster, MedStar Research Institute statistical support.

## APPENDIX

WHI investigators. Program Office: (National Heart, Lung, and Blood Institute, Bethesda, MD) Elizabeth Nabel, Jacques Rossouw, Shari Ludlam, Linda Pottern, Joan McGowan, Leslie Ford, and Nancy Geller. Clinical Coordinating Center: (Fred Hutchinson Cancer Research Center, Seattle, WA) Ross Prentice, Garnet Anderson, Andrea LaCroix, Charles L. Kooperberg, Ruth E. Patterson, Anne McTiernan; (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker; (Medical Research Labs, Highland Heights, KY) Evan Stein; (University of California at San Francisco, San Francisco, CA) Steven Cummings. Clinical Centers: (Albert Einstein College of Medicine, Bronx, NY) Sylvia Wassertheil-Smoller; (Baylor College of Medicine, Houston, TX) Jennifer Hays; (Brigham and Women's Hospital, Harvard Medical School, Boston, MA) JoAnn Manson; (Brown University, Providence, RI) Annlouise R. Assaf; (Emory University, Atlanta, GA) Lawrence Phillips; (Fred Hutchinson Cancer Research Center, Seattle, WA) Shirley Beresford; (George Washington University Medical Center, Washington, DC) Judith Hsia; (Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA) Rowan Chlebowski; (Kaiser Permanente Center for Health Research, Portland, OR) Evelyn Whitlock; (Kaiser Permanente Division of Research, Oakland, CA) Bette Caan; (Medical College of Wisconsin, Milwaukee, WI) Jane Morley Kotchen; (MedStar Research Institute/Howard University, Washington, DC) Barbara V. Howard; (Northwestern University, Chicago/Evanston, IL) Linda Van Horn; (Rush Medical Center, Chicago, IL) Henry Black; (Stanford Prevention Research Center, Stanford, CA) Marcia L. Stefanick; (State University of New York at Stony Brook, Stony Brook, NY) Dorothy Lane; (The Ohio State University, Columbus, OH) Rebecca Jackson; (University of Alabama at Birmingham, Birmingham, AL) Cora E. Lewis; (University of Arizona, Tucson/Phoenix, AZ) Tamsen Bassford; (State University of New York at Buffalo, Buffalo, NY) Jean Wactawski-Wende; (University of California at Davis, Sacramento, CA) John Robbins; (University of California at Irvine, CA) F. Allan Hubbell; (University of California at Los Angeles, Los Angeles, CA) Howard Judd; (University of California at San Diego, LaJolla/Chula Vista, CA) Robert D. Langer; (University of Cincinnati, Cincinnati, OH) Margery Gass; (University of Florida,

Gainesville/Jacksonville, FL) Marian Limacher; (University of Hawaii, Honolulu, HI) David Curb; (University of Iowa, Iowa City/Davenport, IA) Robert Wallace; (University of Massachusetts/Fallon Clinic, Worcester, MA) Judith Ockene; (University of Medicine and Dentistry of New Jersey, Newark, NJ) Norman Lasser; (University of Miami, Miami, FL) Mary Jo O'Sullivan; (University of Minnesota, Minneapolis, MN) Karen Margolis; (University of Nevada, Reno, NV) Robert Brunner; (University of North Carolina, Chapel Hill, NC) Gerardo Heiss; (University of Pittsburgh, Pittsburgh, PA) Lewis Kuller; (University of Tennessee, Memphis, TN) Karen C. Johnson; (University of Texas Health Science Center, San Antonio, TX) Robert Brzyski; (University of Wisconsin, Madison, WI) Gloria E. Sarto; (Wake Forest University School of Medicine, Winston-Salem, NC) Denise Bonds; (Wayne State University School of Medicine/Hutzel Hospital, Detroit, MI) Susan Hendrix.

## REFERENCES

1. Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. *Control Clin Trials* 1998;19:61-109.
2. Rasooly I, Papageorgiou A, Badley E. Comparison of clinical and self-reported diagnosis for rheumatologic outpatients. *Ann Rheum Dis* 1995;54:850-2.
3. Johnson A, Gordon C, Palmer R, Bacon P. The prevalence and incidence of systemic lupus erythematosus in Birmingham, England. *Arthritis Rheum* 1995;38:551-8.
4. Hochberg M, Perlmuter DL, Medsger TA, et al. Prevalence of self-reported physician-diagnosed systemic lupus erythematosus in the USA. *Lupus* 1995;4:454-6.
5. Kvien T, Glennäs A, Knudsrød O, Smedstad L. The validity of self-reported diagnosis of rheumatoid arthritis: results from a population survey followed by clinical examinations. *J Rheumatol* 1996;23:1866-71.
6. Star VL, Scott JC, Sherwin R, Lane N, Nevitt MC, Hochberg MC. Validity of self-reported rheumatoid arthritis in elderly women. *J Rheumatol* 1996;23:1862-5.
7. Barlow JH, Turner AP, Wright CC. Comparison of clinical and self-reported diagnoses for participants on a community-based arthritis self-management programme. *Br J Rheumatol* 1998;37:985-7.
8. Karlson EW, Lee IM, Cook NR, Manson JE, Buring JE, Hennekens CH. Comparison of self-reported diagnosis of connective tissue disease with medical records in female health professionals. *Am J Epidemiol* 1999;150:652-60.
9. Ling SM, Fried LP, Garrett E, Hirsch R, Guralnik JM, Hochberg MC. The accuracy of self-report of physician diagnosed rheumatoid arthritis in moderately to severely disabled older women. *J Rheumatol* 2000;27:1390-4.
10. McAlindon TE, Formica M, Palmer JR, Lafyatis R, Rosenberg L. Assessment of strategies for identifying diagnosed cases of systemic lupus erythematosus through self-report. *Lupus* 2003;12:754-9.
11. Cerhan JR, Saag KG, Merlino LA, Mikuls TR, Criswell LA. Antioxidant micronutrients and risk of rheumatoid arthritis in a cohort of older women. *Am J Epidemiol* 2003;157:345-54.
12. Costenbader KH, Feskanich D, Mandl L, Karlson EW. Smoking intensity, duration, and cessation, and the risk of rheumatoid arthritis in women. *Am J Med* 2006;119:503.e1-9.
13. Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled clinical trial. *JAMA* 2002;288:321-33.
14. The Women's Health Initiative Steering Committee. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: The Women's Health Initiative Randomized Controlled Trial. *JAMA* 2004;291:1701-12.
15. Hays J, Hunt JR, Hubbell A, et al. The Women's Health Initiative recruitment methods and results. *Ann Epidemiol* 2003;13 Suppl:S18-77.
16. Stefanick ML, Cochrane BB, Hsia J, Barad DH, Liu JH, Johnson SR. The Women's Health Initiative postmenopausal hormone trials: overview and baseline characteristics of participants. *Ann Epidemiol* 2003;13 Suppl:S78-86.
17. Karlson EW, Sanchez-Guerrero J, Wright EA, et al. A connective tissue disease screening questionnaire for population studies. *Ann Epidemiol* 1995;5:297-302.
18. Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315-24.
19. Tan E, Cohen A, Fries J, et al. The 1982 American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982;25:1271-7.
20. Altman R, Asch E, Bloch D, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the knee. *Arthritis Rheum* 1986;29:1039-49.
21. Altman R, Alarcon G, Appelrouth, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. *Arthritis Rheum* 1990;33:1601-10.
22. Altman R, Alarcon G, Appelrouth, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1990;34:505-14.
23. Wong A, Harker J, Mittman B, et al. Development and evaluation of a patient self-report case-finding method for rheumatoid arthritis. *Semin Arthritis Rheum* 2004;34:484-99.
24. Liu H, Harker J, Wong A, et al. Case finding for population-based studies of rheumatoid arthritis: comparison of patient self-reported ACR criteria-based algorithms to physician-implicit review for diagnosis of rheumatoid arthritis. *Semin Arthritis Rheum* 2004;33:302-10.