American Journal of Epidemiology
Copyright © 2004 by the Johns Hopkins Bloomberg School of Public Health
All rights reserved

Vol. 160, No. 12
Printed in U.S.A.
DOI: 10.1093/aje/kwh314

Comparison of Self-Report, Hospital Discharge Codes, and Adjudication of Cardiovascular Events in the Women's Health Initiative

Susan R. Heckbert[1], Charles Kooperberg[2,3], Monika M. Safford[4], Bruce M. Psaty[1,5,6], Judith Hsia[7], Anne McTiernan[1,3,5], J. Michael Gaziano[8], William H. Frishman[9], and J. David Curb[10]

[1] Department of Epidemiology, University of Washington, Seattle, WA.
[2] Department of Biostatistics, University of Washington, Seattle, WA.
[3] Fred Hutchinson Cancer Research Center, Seattle, WA.
[4] University of Medicine and Dentistry of New Jersey, Newark, NJ.
[5] Department of Medicine, University of Washington, Seattle, WA.
[6] Department of Health Services, University of Washington, Seattle, WA.
[7] George Washington University School of Medicine, Washington, DC.
[8] Harvard Medical School, Boston, MA.
[9] New York Medical College, Valhalla, NY.
[10] Women's Health Hawaii, Honolulu, HI.

Received for publication April 8, 2004; accepted for publication June 14, 2004.

Limited information is available from large clinical investigations about the agreement among sources of diagnoses for endpoints. The authors used data from the Women's Health Initiative clinical trials and observational study from January 1994 to November 2000 to evaluate the agreement among self-report, hospital discharge codes, and two different levels of physician review of medical records for cardiovascular endpoints. For myocardial infarction, stroke, pulmonary embolism, and venous thrombosis, the agreement of hospital discharge codes or self-report with review by study physicians at clinical centers was substantial (kappa = 0.64–0.84). For coronary revascularization, agreement among these sources of information was substantial to almost perfect (kappa = 0.79–0.92), but for angina, congestive heart failure, and peripheral vascular disease, concordance was only fair to moderate (kappa = 0.37–0.56), indicating that these endpoints remain difficult to classify reliably. Agreement between physician adjudicators at clinical centers and central physician adjudicators was substantial to almost perfect (kappa = 0.67–0.94). The findings also suggest that, for the endpoint of myocardial infarction, physician review of events with hospital discharge codes for angina and congestive heart failure is an important source of validated events, and for stroke, review of all events with cerebrovascular codes is important.

cerebrovascular disorders; classification; clinical trials; coronary disease; epidemiologic methods; longitudinal studies; peripheral vascular diseases; venous thrombosis

Abbreviations: CI, confidence interval; ICD-9, *International Classification of Diseases*, Ninth Revision; WHI, Women's Health Initiative.

In clinical trials and observational studies of moderate size, the number of events often is sufficiently small to allow review of each event by several members of an events committee (1, 2). However, in large multicenter studies involving tens of thousands of participants and multiple endpoints, streamlined methods of endpoint ascertainment are necessary (3–5). In some studies, events are adjudicated by study physicians at clinical centers, and all or a portion of events is reviewed by a central adjudicator for quality control purposes (6, 7). Nonetheless, limited information is available from such studies, particularly from clinical trials, about agreement among sources of diagnoses for cardiovascular events.

The Women's Health Initiative (WHI), a complex clinical investigation of strategies for long-term prevention of common diseases in over 160,000 postmenopausal women at 40 clinical centers, includes a set of clinical trials and an observational study (8). In the WHI, information on cardio-

Reprint requests to Dr. Susan R. Heckbert, University of Washington Cardiovascular Health Research Unit, 1730 Minor Avenue, Suite 1360, Seattle, WA 98101-1448 (e-mail: heckbert@u.washington.edu).

vascular events is collected from four sources: participant self-report, hospital discharge codes (*International Classification of Diseases*, Ninth Revision (ICD-9), disease and procedure codes), review of medical records by a physician at the participant's clinical center (local adjudicator, generally one or two physicians at each of 40 centers), and for selected endpoints, review of medical records by one of the 8–10 physician members of the WHI Central Cardiovascular Review Committee (central adjudicator). This setting provides a unique opportunity to evaluate the source of diagnosis information and extent of agreement for cardiovascular endpoints using various methods of ascertainment.

MATERIALS AND METHODS

Setting

The WHI clinical trials, including a total of 68,132 women enrolled during 1993–1998, test whether long-term preventive measures including two distinct regimens of hormone replacement therapy, low-fat diet, and supplementation with calcium and vitamin D decrease the incidence of cardiovascular disease, certain cancers, and fractures. The observational study follows 93,676 postmenopausal women enrolled during 1994–1998 who were ineligible for, or declined enrollment in, the clinical trials (8). The estrogen-plus-progestin clinical trial was stopped early in July 2002 at the recommendation of the data and safety monitoring board, because the overall risks exceeded the benefits (9), and the estrogen-only trial was stopped early in March 2004 (10). The other clinical trials and the observational study continue.

The primary endpoint for the hormone replacement trials was nonfatal myocardial infarction plus coronary heart disease death, and secondary endpoints included hospitalized stroke, angina, coronary revascularization, congestive heart failure, transient ischemic attack, and peripheral vascular disease. While all of these cardiovascular endpoints are also ascertained for all participants in both other clinical trials and the observational study, deep venous thrombosis and pulmonary embolism were safety endpoints ascertained only for the hormone clinical trials.

For the hormone clinical trials, myocardial infarction, stroke, pulmonary embolism, and deep venous thrombosis received both local and central adjudication. In case of disagreement, the WHI protocol calls for review by a second central adjudicator and, if needed, review by committee. The final centrally adjudicated diagnosis for cardiovascular disease reported in study publications for the hormone trials is the diagnosis reached at the end of this review process. For the clinical trials evaluating a low-fat diet and supplementation with calcium and vitamin D, cardiovascular disease is not the primary endpoint. According to WHI protocol, publications reporting results for these two trials will report the centrally adjudicated diagnoses for their primary endpoints (cancers and selected fractures) and the locally adjudicated diagnoses for cardiovascular endpoints. All clinical trial components report the centrally adjudicated cause of death. In observational study publications, the locally adjudicated diagnosis is reported for both cardiovascular endpoints and fatal events. Nonetheless, during the early years of the study covered by this analysis, myocardial infarction, angina, congestive heart failure, coronary bypass, and angioplasty received both local and central adjudication for both the clinical trials and the observational study, permitting an evaluation of the extent of agreement across all study components.

Data collection and adjudication

Informed consent was obtained from participants, and the institutional review boards of the study centers approved the study protocol. Participants or their proxies completed standardized questionnaires, every 6 months for clinical trial participants and every year for observational study participants, that asked about all hospitalizations, outpatient coronary revascularization, and outpatient diagnosis of deep venous thrombosis since the last update. Potential events were identified from these participant self-reports of hospitalizations and selected outpatient diagnoses. For a specified list of common procedures and conditions not of interest to the study, no hospital records were requested. For all other events, medical records were requested from the hospital or outpatient facility and death certificates were obtained from the state or the family, and these records were reviewed by the local adjudicator. Elements of the medical record requested included hospital discharge codes, the discharge summary, laboratory studies, electrocardiograms, diagnostic test reports, and procedure reports. Hospital discharge codes were provided for 63 percent of events for which medical records were requested and received.

The local adjudicator was not the participant's personal physician and, for clinical trial participants, was blinded to randomization assignment. For all potential events evaluated in this analysis (except stroke and peripheral vascular disease events), if the local adjudicator identified a cardiovascular event meeting WHI study definitions, the medical records were forwarded to a WHI central adjudicator. Central adjudicators are members of the WHI Central Cardiovascular Review Committee, which establishes procedures for adjudication and meets periodically to review difficult cases. In addition, for this analysis, the interrater agreement between central adjudicators was assessed in a subset of 81 cardiovascular cases that were reviewed by two central adjudicators. The WHI criteria for cardiovascular endpoints were adapted from standardized criteria and are reported elsewhere (11).

Statistical analysis

Included in this analysis were all participant reports for any WHI endpoint, including noncardiovascular endpoints, from both the clinical trials and the observational study, on which processing was complete between January 1, 1994, and November 30, 2000. There was a lag between the event date and the close of processing, and the sample included only reports on which processing was complete. Cardiovascular events of interest included first episodes after study enrollment of fatal or nonfatal myocardial infarction, angina, congestive heart failure, occurrences of coronary artery bypass graft and coronary angioplasty, stroke, transient ischemic attack, peripheral vascular disease, and, for the

hormone trials only, the safety endpoints of pulmonary embolism or deep venous thrombosis. A particular woman could appear more than once in the analysis if, for example, she reported a first myocardial infarction as well as a first stroke during the period of analysis. The proportion of women reporting a history of myocardial infarction at enrollment was 2.1 percent in the clinical trials and 2.5 percent in the observational study; for stroke the comparable numbers were 1.1 percent and 1.5 percent, and for deep venous thrombosis or pulmonary embolism they were 3.2 percent and 4.4 percent.

We evaluated positive predictive value (proportion of positives reported by the test method that were verified by the reference method), sensitivity (proportion of positives reported by the reference method that were also reported as positives by the test method), and specificity (proportion of negatives reported by the reference method that were also reported as negatives by the test method) as measures of validity. We used the kappa statistic with exact binomial 95 percent confidence intervals as a measurement of the extent of agreement beyond chance alone (12). Kappa values of 0.81–0.99 were considered to represent almost perfect agreement and of 0.61–0.80 substantial, 0.41–0.60 moderate, and 0.21–0.40 fair agreement (13). Because of the relative rarity of cardiovascular endpoints among all the potential events reported by participants, in analyses that assessed agreement of hospital discharge codes and self-report with local adjudication, specificity was over 99 percent with very narrow 95 percent confidence intervals, except as noted in the text.

RESULTS

Comparison of hospital discharge codes with local adjudication

Between January 1, 1994, and November 30, 2000, there were 34,016 participant reports of events on which processing was complete and where hospital discharge codes (ICD-9 disease and/or procedure codes) were requested and available. The agreement between selected ICD-9 codes and local adjudicator diagnosis was examined, with the local adjudicator opinion used as the reference method (table 1). The proportion of events with an ICD-9 code indicating myocardial infarction (code 410) or cardiac arrest (codes 427.4 and 427.5) that were verified as myocardial infarction by the local adjudicator (the positive predictive value) was 78 percent. The sensitivity of the selected ICD-9 codes for myocardial infarction overall was 80 percent, and the specificity was 99.4 (95 percent confidence interval (CI): 99.3, 99.5) percent. The kappa statistic was 0.78. There were 169 events identified by the local adjudicator as myocardial infarction that did not have a code of 410 or 427.4–427.5; most of these events had a code indicating angina ($n = 132$), congestive heart failure ($n = 10$), or other cardiovascular disease ($n = 11$). Agreement was also substantial for stroke and for the safety endpoints of pulmonary embolism and deep venous thrombosis. The 127 events identified by the local adjudicator as a stroke that did not have one of the listed stroke codes were coded as transient ischemic attack (code 435; $n = 23$) or as another cerebrovascular disease (code 433, 437, or 438; $n = 58$), or they had no cerebrovascular disease code ($n = 46$). For angina, congestive heart failure, and peripheral vascular disease, the positive predictive value and kappa statistics were considerably lower than for myocardial infarction or stroke, and the specificities were 93, 98, and 99 percent, respectively. By contrast, all indices of agreement were quite high for both types of coronary revascularization.

Comparison of self-report with local adjudication

Concordance between self-reported cardiovascular endpoints and local adjudicator diagnosis, used as the reference method, was compared for all 99,500 participant reports on which processing was complete, regardless of whether hospitalization records were requested or hospital discharge codes received (table 2). For self-reports of cardiovascular disease, hospitalization records were always requested. Stroke and transient ischemic attack are combined because participants were asked a single question about both conditions. The proportion of self-reported events verified by the local adjudicator (positive predictive value) was 58–75 percent for myocardial infarction, congestive heart failure, angioplasty, stroke/transient ischemic attack, peripheral vascular disease, and deep venous thrombosis, but it was considerably lower (37 percent) for angina. By contrast, self-reports of coronary bypass surgery and pulmonary embolism had high rates of verification by local adjudicators (positive predictive values: 90 and 86 percent, respectively). The kappa statistics for agreement between self-report and local adjudication for angina, congestive heart failure, and peripheral vascular disease were 0.37, 0.48, and 0.53, respectively.

Source of diagnosis for events identified by local adjudication

Among the 34,016 participant reports of events on which processing was complete and where hospital discharge codes were requested and available during the study period, there were 11,272 participant reports where the self-report, the hospital discharge code, or the local adjudicator identified a cardiovascular event. Considering local adjudication as the reference method, the source of the diagnosis for various cardiovascular endpoints is shown in table 3. For example, for myocardial infarction, of the 847 events judged to meet WHI criteria by the local adjudicator, 55 percent were identified by all three sources: self-report, hospital discharge code, and the local adjudicator; 6 percent by self-report and the local adjudicator; 25 percent by hospital discharge code and the local adjudicator; and 14 percent by the local adjudicator only (an event was identified during a hospitalization without a hospital discharge code for myocardial infarction, and a myocardial infarction was not reported by the participant). A total of 39 percent of myocardial infarction events identified by the local adjudicator were not identified by self-report. For angina and congestive heart failure, around 60 percent of the events identified by the local adjudicators were not identified by self-report, and for peripheral vascular disease, the number was 48 percent. Conversely, there were large numbers of events with hospital discharge codes indi-

TABLE 1. Agreement between selected ICD-9* codes and local adjudicator diagnosis for 34,016 reports of potential events of all types reviewed by local adjudicators, Women's Health Initiative, January 1994–November 2000

Endpoint	ICD-9† code	Code = yes, local = yes‡ (no.)	Code = yes, local = no (no.)	Code = no, local = yes (no.)	Code = no, local = no (no.)	Positive predictive value§		Sensitivity§		Kappa statistic	95% CI*
						%	95% CI	%	95% CI		
Myocardial infarction	Overall	678	195	169	32,974	78	75, 80	80	77, 83	0.78	0.76, 0.80
	410	670	137								
	427.4	4	30								
	427.5	4	28								
Angina	Overall	1,533	2,281	176	30,026	40	39, 42	90	88, 91	0.52	0.50, 0.54
	411	897	320								
	413	284	366								
	414	352	1,595								
Congestive heart failure	Overall	628	757	167	32,464	45	43, 48	79	76, 82	0.56	0.53, 0.59
	428	603	638								
	425	25	119								
Coronary bypass	36.xx¶	576	29	71	33,340	95	93, 97	89	86, 91	0.92	0.90, 0.94
Angioplasty	36.xx#	805	48	132	33,031	94	93, 96	86	84, 88	0.90	0.88, 0.91
Stroke	Overall	582	138	127	33,169	81	78, 84	82	79, 85	0.81	0.79, 0.83
	430	31	11								
	431	76	6								
	432.0–432.1	5	16								
	432.9	3	2								
	434	375	64								
	436	92	39								
Transient ischemic attack	435	319	122	119	33,456	72	68, 76	73	68, 77	0.72	0.69, 0.76
Peripheral vascular disease	Overall	121	271	78	33,546	31	26, 36	61	54, 68	0.40	0.35, 0.46
	440.2	94	80								
	443.9	27	191								
Pulmonary embolism**	415.1	50	8	11	6,765	86	75, 94	82	70, 91	0.84	0.77, 0.91
Deep venous thrombosis**	Overall	104	35	17	6,678	75	67, 82	86	78, 92	0.80	0.74, 0.85
	451.x††	14	8								
	453.x‡‡	90	27								

* ICD-9, *International Classification of Diseases*, Ninth Revision; CI, confidence interval.
† ICD-9 categories are mutually exclusive and hierarchical in the order listed.
‡ Code = ICD-9 code; local = local adjudicator diagnosis.
§ Local adjudicator diagnosis was used as the reference method.
¶ Procedure codes for coronary bypass surgery are 36.xx, where xx = any of the following two digits after the decimal: 14, 13, 12, 11, 10, 16, 15, 19, 3, 2, 17, 03.
# Procedure codes for angioplasty are 36.xx, where xx = any of the following two digits after the decimal: 05, 02, 06, 01, 09.
** Pulmonary embolism and deep venous thrombosis received local adjudicator review only for women in the hormone trials (6,834 reports of potential events of all types).
†† ICD-9 codes for deep venous thrombosis are 451.x, where x = any of the following digits after the decimal: 1, 2, 8, 9.
‡‡ ICD-9 codes for deep venous thrombosis are 453.x, where x = any of the following digits after the decimal: 0, 1, 2, 8, 9.

cating angina, congestive heart failure, and peripheral vascular disease that were determined not to meet WHI criteria for these cardiovascular events by local adjudicators.

Comparison of local and central adjudication

Local adjudication and the initial central adjudication were complete on 3,634 cardiac events and 144 pulmonary embolism or venous thrombosis events identified by the local adjudicators (table 4). Peripheral vascular disease is not centrally adjudicated in WHI, and during the period covered by this analysis, stroke/transient ischemic attack was not yet centrally adjudicated; therefore these endpoints do not appear in table 4. The agreement between local and central adjudication as measured by the kappa statistic was 0.79–0.94 for myocardial infarction, congestive heart failure, both types of revascularization, and the safety endpoints of pulmonary embolism and venous thrombosis, but it was 0.67 for angina. Central adjudication identified fewer events of all types than did local adjudication. To assess the interrater agreement between central adjudicators, two separate central adjudicators conducted reviews in a subset of 81 cardiac

TABLE 2. Agreement between self-reported cardiovascular outcomes and local adjudicator diagnosis for 99,500 participant reports of potential events, Women's Health Initiative, January 1994–November 2000

Endpoint	Self-report = yes, local = yes* (no.)	Self-report = yes, local = no (no.)	Self-report = no, local = yes (no.)	Self-report = no, local = no (no.)	Positive predictive value†		Sensitivity†		Kappa statistic	95% CI‡
					%	95% CI	%	95% CI		
Myocardial infarction	633	296	404	98,167	68	65, 71	61	58, 64	0.64	0.61, 0.67
Angina	847	1,458	1,226	95,969	37	35, 39	41	39, 43	0.37	0.35, 0.39
Congestive heart failure	372	203	597	98,328	65	61, 69	38	35, 42	0.48	0.45, 0.51
Coronary bypass	691	80	80	98,649	90	87, 92	90	87, 92	0.90	0.87, 0.92
Angioplasty	953	323	187	98,037	75	72, 77	84	81, 86	0.79	0.77, 0.81
Stroke/transient ischemic attack	1,164	444	265	97,627	72	70, 75	81	79, 83	0.76	0.74, 0.78
Peripheral vascular disease	129	93	133	99,145	58	51, 65	49	43, 55	0.53	0.48, 0.59
Pulmonary embolism	62	10	15	99,413	86	76, 93	81	70, 89	0.83	0.77, 0.90
Deep venous thrombosis	112	55	32	99,301	67	59, 74	78	70, 84	0.72	0.66, 0.78

* Local = local adjudicator diagnosis.
† Local adjudicator diagnosis was used as the reference method.
‡ CI, confidence interval.

events; the kappa statistics were 0.81 for myocardial infarction, 0.65 for angina, 0.71 for congestive heart failure, 0.92 for coronary bypass surgery, and 0.81 for angioplasty.

Local and central adjudications were compared for 1,518 deaths. With the initial central adjudication as the reference method, for the endpoint of death due to definite or possible coronary heart disease, the positive predictive value was 80 (95 percent CI: 74, 85) percent, and the kappa statistic was 0.75 (95 percent CI: 0.70, 0.80). With the final central adjudication decision as the reference method for 634 deaths among participants in the clinical trials, the positive predictive value for local adjudication was 96 (95 percent CI: 90, 99) percent, and the kappa statistic was 0.89 (95 percent CI: 0.84, 0.94). For deaths due to cerebrovascular disease in both the clinical trials and the observational study, with the initial central adjudication as the reference method, the positive predictive value was 86 (95 percent CI: 78, 92) percent, and the kappa statistic was 0.84 (95 percent CI: 0.79, 0.90).

DISCUSSION

In the WHI, the use of hospital discharge codes or self-report to identify cardiovascular events without local adjudication resulted in some misclassification of myocardial infarction, stroke, pulmonary embolism, and venous thrombosis, as well as considerable misclassification of angina, congestive heart failure, and peripheral vascular disease. On the other hand, the use of procedure codes to identify coronary revascularization procedures, particularly coronary bypass surgery, appeared to be reasonably accurate and complete.

TABLE 3. Source of diagnoses for 11,272 potential events identified as cardiovascular events by self-report, hospital discharge codes, or local adjudication, Women's Health Initiative, January 1994–November 2000

Endpoint	Identified as cardiovascular events by local adjudicators									Determined not to be cardiovascular events by local adjudicators			
	Total (no.)	Self-report, code, and local = yes*		Self-report and local = yes		Code and local = yes		Local = yes		Total (no.)	Self-report and code = yes (no.)	Code = yes (no.)	Self-report = yes (no.)
		No.	%†	No.	%†	No.	%†	No.	%†				
Myocardial infarction	847	466	55	48	6	212	25	121	14	314	41	154	119
Angina	1,709	643	38	55	3	890	52	121	7	2,430	382	1,899	149
Congestive heart failure	795	276	35	28	4	352	44	139	18	825	58	699	68
Coronary bypass	647	527	82	56	9	49	8	15	2	137	20	9	108
Angioplasty	937	695	74	97	10	110	12	35	4	209	32	16	161
Stroke	709	517	73	97	14	65	9	30	4	314	64	74	176
Transient ischemic attack	438	241	55	69	16	78	18	50	11	210	66	56	88
Peripheral vascular disease	199	82	41	22	11	39	20	56	28	296	5	266	25
Pulmonary embolism	61	40	66	8	13	10	16	3	5	22	0	8	14
Deep venous thrombosis	121	83	69	12	10	21	17	5	4	52	9	26	17

* Code = *International Classification of Diseases*, Ninth Revision, code; local = local adjudicator diagnosis.
† Percentage of the total number of events judged to meet Women's Health Initiative criteria by local adjudicators.

TABLE 4. Agreement between local adjudicator diagnosis and central adjudicator diagnosis for 3,634 cardiac events and 144 pulmonary embolism or venous thrombosis events reviewed by central adjudicators, Women's Health Initiative, January 1994–November 2000

Endpoint	Local = yes, central = yes* (no.)	Local = yes, central = no (no.)	Local = no, central = yes (no.)	Local = no, central = no (no.)	Sensitivity†		Specificity†		Kappa statistic	95% CI‡
					%	95% CI	%	95% CI		
Myocardial infarction	558	131	93	2,852	86	83, 88	96	95, 96	0.80	0.77, 0.82
Angina	1,047	381	181	2,025	85	83, 87	84	83, 86	0.67	0.64, 0.70
Congestive heart failure	481	151	53	2,949	90	87, 92	95	94, 96	0.79	0.76, 0.82
Coronary bypass	475	35	20	3,104	96	94, 98	99	98, 99	0.94	0.92, 0.96
Angioplasty	662	74	48	2,850	93	91, 95	97	97, 98	0.89	0.88, 0.91
Pulmonary embolism	48	8	1	87	98	89, 100	92	84, 96	0.87	0.78, 0.95
Deep venous thrombosis	89	12	1	42	99	94, 100	78	64, 88	0.80	0.70, 0.90

* Local = local adjudicator diagnosis; central = central adjudicator diagnosis.
† Central adjudicator diagnosis was used as the reference method.
‡ CI, confidence interval.

Agreement was very good to excellent between local and central adjudicators for all cardiovascular endpoints studied. Concordance was best for myocardial infarction, congestive heart failure, coronary bypass surgery, angioplasty, pulmonary embolism, and venous thrombosis but was less substantial for angina. For the underlying cause of death, agreement between local and central adjudicators was very good to excellent for deaths due to combined definite or possible coronary disease and for deaths due to cerebrovascular disease.

Strengths of this analysis include the large numbers of outcomes investigated, the extensive and detailed follow-up of participants, the completeness of the data, the representation of participants from a wide variety of US communities, and the ability to report the source of diagnosis for events verified by local adjudication. The WHI included relatively few women with low socioeconomic status or from rural areas, and it did not include men. Thus, the results reported here might not be generalizeable to these groups. Hospital discharge codes were available for only 63 percent of self-reported events for which medical records were received. Thus, conclusions about agreement of hospital discharge codes with local adjudicator diagnosis might be biased if missing discharge code data did not occur at random. In this study, nonfatal cardiovascular events other than venous thrombosis and angioplasty were limited to those resulting in hospitalization. Thus, our analysis did not include outpatient events or events that were not clinically recognized. For the assessment of agreement between local and central adjudication, only those events identified by local adjudicators were forwarded for further review by central adjudicators during the period of this analysis. Self-reported events found not to meet WHI criteria for events by local adjudicators were not forwarded for central adjudication. This screening process could have affected the estimates of validity and agreement for the comparison of local and central adjudication. The number of pulmonary embolism and venous thrombosis events was small because they were ascertained in only the hormone therapy trials, and findings for these safety endpoints should be interpreted with caution. Finally, information about the circumstances of death in WHI was generally limited to the death certificate and records from the most recent hospitalization during the study; informant interviews were available for only a small proportion of fatal cases. Using information on the circumstances of death from medical records and interviews with a variety of informants, including physicians and medical examiners, the Atherosclerosis Risk in Communities Study documented that coronary disease is listed as the cause of death on death certificates more often than is supported by the evidence (14), a finding that has been observed in other studies (15).

The extent of agreement between hospital discharge codes and physician adjudicator review for various cardiovascular endpoints was similar in the WHI and the Cardiovascular Health Study (1), both of which used self-report as the starting point for investigating events and used very similar criteria for defining cardiovascular events. Several other studies have reported the extent of agreement between hospital discharge codes and physician review for the endpoint of myocardial infarction (4, 16–18). Although the positive predictive value of ICD-9 code 410 for myocardial infarction was uniformly quite high in these studies (range: 79–90 percent), the sensitivity varied considerably (range: 36–89 percent), possibly related to differences in the range of ICD-9 codes investigated and varying criteria for myocardial infarction. As in the First National Health and Nutrition Examination Survey Epidemiologic Follow-up Study, the number of coronary disease events identified in the WHI increased as more sources of data were used (19). Across all hospital discharge codes corresponding to stroke, the positive predictive value was 90 percent in the Cardiovascular Health Study and 81 percent in the WHI. Similar results have been reported for stroke from Saskatchewan (20).

For the two hormone clinical trials in WHI, the final centrally adjudicated diagnosis for cardiovascular endpoints reported in study publications is the diagnosis reached after the initial central review and any additional central review required because of disagreement with the local adjudicator. Since a fraction of central reviews initially discordant with the local adjudication are overturned by subsequent central reviews, the extent of agreement with local adjudication is higher for the final central diagnosis than for the initial

central diagnosis, as was observed in this analysis for deaths due to coronary disease.

In the WHI estrogen-plus-progestin trial, the reported positive predictive value for local adjudication as compared with central adjudication was 90 percent for myocardial infarction and 97 percent for death due to definite or possible coronary disease (21). These rates compare favorably with those from other large multicenter trials, such as the Antihypertensive and Lipid Lowering Treatment to Prevent Heart Attack Trial (ALLHAT), in which the reported positive predictive value for local adjudication as compared with central adjudication was 92 percent (72 of 78 events; 95 percent CI: 84, 97) for the composite endpoint of nonfatal myocardial infarction and fatal coronary heart disease (6).

In very large studies, cost-effective adjudication of cardiovascular endpoints is a challenge. Findings from the WHI suggest that, for the endpoint of myocardial infarction, physician review of events with hospital discharge codes for angina and congestive heart failure is an important source of validated events, and for stroke, review of all events with cerebrovascular codes is important. For coronary bypass, agreement was excellent for both self-report and hospital discharge codes, with physician review of medical records suggesting that any of these sources of information could be used; however, for angioplasty, agreement of self-report with local adjudication was slightly lower than for coronary bypass. For the clinical syndromes of angina and congestive heart failure and for peripheral vascular disease, concordance among the various sources of information was only fair to good, suggesting that these endpoints remain difficult to classify reliably. These findings may prove useful to investigators planning large-scale investigations with cardiovascular endpoints.

ACKNOWLEDGMENTS

The Women's Health Initiative program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services.

REFERENCES

1. Ives DG, Fitzpatrick AL, Bild DE, et al. Surveillance and ascertainment of cardiovascular events. The Cardiovascular Health Study. Ann Epidemiol 1995;5:275–85.
2. The SHEP Cooperative Research Group. Rationale and design of a randomized clinical trial on prevention of stroke in isolated systolic hypertension. J Clin Epidemiol 1988;41:1197–208.
3. Davis BR, Cutler JA, Gordon DJ, et al. Rationale and design for the Antihypertensive and Lipid Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). ALLHAT Research Group. Am J Hypertens 1996;9:342–60.
4. White AD, Folsom AR, Chambless LE, et al. Community surveillance of coronary heart disease in the Atherosclerosis Risk in Communities (ARIC) Study: methods and initial two years' experience. J Clin Epidemiol 1996;49:223–33.
5. The HOPE Study Investigators. The HOPE (Heart Outcomes Prevention Evaluation) Study: the design of a large, simple randomized trial of an angiotensin-converting enzyme inhibitor (ramipril) and vitamin E in patients at high risk of cardiovascular events. Can J Cardiol 1996;12:127–37.
6. The ALLHAT Collaborative Research Group. Major cardiovascular events in hypertensive patients randomized to doxazosin vs chlorthalidone: the Antihypertensive and Lipid Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). JAMA 2000;283:1967–75.
7. Lauer MS, Topol EJ. Clinical trials—multiple treatments, multiple end points, and multiple lessons. JAMA 2003;289:2575–7.
8. The 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.
9. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. JAMA 2002;288:321–33.
10. Anderson GL, Limacher M, Assaf AR, et al. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. JAMA 2004;291:1701–12.
11. Curb D, McTiernan A, Heckbert SR, et al. Outcomes ascertainment and adjudication methods in the Women's Health Initiative. Ann Epidemiol 2003;13(suppl):S122–8.
12. Pepe MS. The statistical evaluation of medical tests for classification and prediction. Oxford, United Kingdom: Oxford University Press, 2003.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
14. Coady SA, Sorlie PD, Cooper LS, et al. Validation of death certificate diagnosis for coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) Study. J Clin Epidemiol 2001;54:40–50.
15. Lauer MS, Blackstone EH, Young JB, et al. Cause of death in clinical research: time for a reassessment? J Am Coll Cardiol 1999;34:618–20.
16. Mahonen M, Salomaa V, Brommels M, et al. The validity of hospital discharge register data on coronary heart disease in Finland. Eur J Epidemiol 1997;13:403–15.
17. Austin PC, Daly PA, Tu JV. A multicenter study of the coding accuracy of hospital discharge administrative data for patients admitted to cardiac care units in Ontario. Am Heart J 2002;144:290–6.
18. Pladevall M, Goff DC, Nichaman MZ, et al. An assessment of the validity of ICD code 410 to identify hospital admissions for myocardial infarction: the Corpus Christi Heart Project. Int J Epidemiol 1996;25:948–52.
19. Madans JH, Reuben CA, Rothwell ST, et al. Difference in morbidity measures and risk factor identification using multiple data sources: the case of coronary heart disease. Stat Med 1995;14:643–53.
20. Liu L, Reeder B, Shuaib A, et al. Validity of stroke diagnosis on hospital discharge records in Saskatchewan, Canada: implications for stroke surveillance. Cerebrovasc Dis 1999;9:224–30.
21. Manson JE, Hsia J, Johnson KC, et al. Estrogen plus progestin and the risk of coronary heart disease. N Engl J Med 2003;349:523–34.