Use of administrative data to increase the practicality of clinical trials: Insights from the Women's Health Initiative

Garnet L Anderson[1,*], **Carolyn J Burns**[1], **Joseph Larsen**[1], and **Pamela A Shaw**[2]

[1]WHI Clinical Coordinating Center, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[2]Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA, USA

Abstract

Background—To reduce research costs in the context of pragmatic trials, consideration is given to using administrative data (Medicare claims) to ascertain clinical outcomes.

Methods—In the historical context of the Women's Health Initiative, the correspondence between selected cardiovascular events derived from Medicare claims were compared to those documented and adjudicated in this large-scale prevention trial.

Results—Classification performance varies somewhat by type of outcome but hazard ratios and confidence intervals derived from the two data sources were quite comparable.

Conclusions—These encouraging results provided the needed support to launch a new embedded pragmatic trial of physical activity that will rely heavily on Medicare claims to ascertain cardiovascular disease incidence in the majority of those randomized.

Keywords

Pragmatic trials; mismeasured outcomes; misclassification; claims data; Medicare linkage

Introduction

Pragmatic trials are intended to provide strong, reliable data to inform clinical practice. Their designs are typically streamlined to answer an overall question of treatment impact in a usual practice setting--in contrast to the intense subject monitoring and extensive data collection often required for mechanistic trials. Optimally, pragmatic trials test scalable interventions to support direct generalization to the target population and do so within a current health care context to minimize cost.

Research costs per participant are usually dramatically lower in pragmatic trials but the large sample sizes required to detect clinical benefit means they are rarely inexpensive. Furthermore, information on the broader effects of an intervention, including off-target and unanticipated effects has implications for translation to practice. These are most reliably

assessed as outcomes in the context of a randomized trial. To properly capture the expected and unexpected effects of treatment, the scope of the trial could expand considerably, driving up research costs.

The demand for improving generalizability and broadening data collection while reducing costs of clinical trials have led to an interest in re-purposing data collected for other reasons. Administrative data, such as health insurance or Medicare claims data, have long been used in health services research, but they are not commonly used to capture clinical outcomes data in clinical trials. To understand how we might meet these demands, and particularly to conduct other, more efficient randomized trials within an existing study cohort, investigators in the Women's Health Initiative (WHI) program have been evaluating the use of Medicare claims data for outcomes assessment. The results have been sufficiently promising that a pragmatic trial of physical activity has recently been funded by the National Heart Lung and Blood Institute that will rely heavily upon Medicare data. Here we briefly describe both WHI and the newly embedded trial, called WHI Strong and Healthy (WHISH), and review the WHI evaluation of Medicare data that supported the WHISH pragmatic trial design.

Women's Health Initiative overview

The Women's Health Initiative was launched in 1993 to examine the effects of menopausal hormone therapy, a low-fat diet, and calcium and vitamin D supplements on risk of major chronic diseases in postmenopausal women. Overall, 68,132 women aged 50-79 years participated in one, two or all three of these randomized controlled trials. Women who were excluded from the trials for safety, adherence, or competing risk criteria or who were not interested were offered enrollment in the parallel observational study (n=93,676). Study design and implementation details have been published.[1,2]

The WHI trials were pragmatic in the sense that eligibility was broadly defined, recruitment was population-based, interventions were delivered in a cost-effective manner, and there was a broad assessment of health effects including all hospitalized events.[1] Furthermore, the results of each trial were intended to directly inform clinical practice. WHI was not a so-called large simple trial, however, in that the entire infrastructure was created to purpose. The data collected, including adherence assessments, selected health updates, medical records retrieval and adjudication of all key health events were developed to support this study--a costly, burdensome activity.[2,3]

In 2005 after all of the trials ended, participant follow-up was extended without further intervention for all re-consenting participants for the purpose of documenting longer term or late effects of the interventions. The frequency of health updates was reduced to annual self-reports but medical record collection and adjudication continued using the same mechanisms. To accommodate reduced funding beginning in 2010, adjudication of clinical events was limited to approximately 25% of the continuing participants. This subset, referred to as the Medical Records Cohort, was composed of former menopausal hormone trial participants and all African-American and Hispanic participants, the subset for which additional data was expected to be most informative. The remaining participants still in active follow-up, called the Self-Report Cohort, received the same annual health update

questionnaires but their medical records were no longer collected nor health outcomes adjudicated.

Throughout follow-up, outcomes data have been supplemented with death information obtained through linkage to the National Death Index. Subsequent linkage to Medicare data from the Center for Medicare and Medicaid Services (referred to subsequently as Medicare) was originally pursued to support analyses of an even broader range of health conditions than were captured through the active follow-up mechanisms.

WHI Strong and Healthy (WHISH) overview

The WHISH trial, recently funded by the National Heart Lung and Blood Institute (Drs. Marcia Stefanick, Charles Kooperberg, and Andrea LaCroix, Principal Investigators), is a pragmatic randomized trial embedded within the WHI cohort, designed to determine whether a public health-based intervention to increase physical activity would reduce cardiovascular disease risk in older women.[4] The intervention being tested (Go4Life) is available to the general public through the National Institutes on Aging.[5] Because WHI participants still in active follow-up have consented to be in research, and have agreed to provide ongoing updates on their health status and to allow linkage to national data sources, the necessary outcomes data for WHISH are already available through WHI. Taking advantage of this existing infrastructure, the trial uses Zelen's randomized consent design[6] with an opt-out strategy for consenting participants selected for intervention. Cardiovascular health outcomes for women in the Medical Records Cohort will be based on traditional physician adjudication but those for women in the Self-Report Cohort will be obtained largely through linkage to Medicare and the National Death Index.

The primary analysis will be based on a Cox regression model under the intention-to-treat principle, comparing those randomized to intervention, regardless of their consent status to the intervention, to those randomized to control with stratification by cohort membership (Medical Records vs Self-Report). The validity of the comparisons in the Self-Report Cohort relies primarily on the retrospective evaluation of these data within the original WHI study.

Access to Medicare data

The WHI Clinical Coordinating Center sought Medicare files dating back to 1991 to link to the WHI participant database. Medicare routinely provides a data file documenting the coverage status of enrollees over time (Denominator file) as well as files organized around reimbursement claim types: Inpatient (MedPar), Outpatient, Home Health, Skilled Nursing, Hospice, Durable Medical Equipment, Carrier, and Part D—Prescription Drug.

With Institutional Review Board approval, WHI provided Medicare with personal identifiers for 151,169 WHI participants, the entire cohort for whom valid social security numbers were available. Of these, 150,792 had reached age-eligibility for Medicare by the time of latest linkage (2012). Medicare returned a linkage file with 147,998 records matched by social security numbers who were Medicare enrollees. Of these, 1161 were dropped because they did not appear in the Medicare Denominator file between 1991 and 2012 or because Medicare indicated the enrollee was male. Of the remaining 146,837 linked records based on

social security number, 143,058 had exact matches for either date of birth or death. An additional 2633 were viewed as acceptable linkages after allowing for small discrepancies in birth date but requiring a match on 5-digit zip-code. Overall, linkage was considered adequately established for 145,691 participant records, 96.6% of age-eligible participants with a valid social security number.

Although the Medicare linkage was highly successful, the availability of claims data is limited by the inclusion of study participants enrolled in health maintenance/managed care organizations (HMOs) which may vary over time. HMOs function under a capitated system where Medicare does not process individual claims for visits, procedures, etc. Rather providers receive reimbursements for their Medicare patients based on an overall cost per enrollee. Thus while an HMO enrollee may be found indicating the study participant was covered by Medicare, details of their provider visits are not usually available.

Medicare coverage for WHI participants increased with the aging of the cohort to nearly 80% (Figure 1). Roughly one-third of these Medicare enrollees received their health care through HMOs in any calendar year, limiting the availability of their claims data. Because individuals have regular opportunities to change type of coverage, additional gaps in the administrative data stream can occur, although these appear to be relatively infrequent. For example, between January 1, 2011 and January 1, 2012, 2.12% of WHI participants in Medicare changed types of coverage, with 0.55% switching from HMO to fee-for-service and a larger fraction (1.57%) migrating from fee-for-service to HMO coverage.

Agreement rates between outcomes ascertained through WHI and Medicare

The validity of cardiovascular health events ascertained by algorithms applied to Medicare data was assessed by comparing events in the same women found by the current standard of protocol-driven data collection and physician adjudication of medical records. These analyses were restricted to the Inpatient Medicare data to evaluate it as a potential alternate source of data for the adjudicated hospitalized events collected throughout WHI follow-up. They were further restricted to those study subjects who were continuously enrolled in fee-for-service Medicare (Part A) during their WHI participation. The continuous enrollment helps to ensure that any discrepancies would not be attributable to changing coverage. Previously published analyses reviewed here examined selected cardiovascular diagnoses (acute myocardial infarction), stroke and abdominal aortic aneurysm) and procedures (coronary revascularization, peripheral vascular procedures of the lower extremity, and carotid artery treatments).[7-9]

Algorithms using International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis and procedure codes were taken from the Chronic Conditions Warehouse.[10] The algorithms themselves are not complex but numerous codes may be required to capture a condition (See codes for selected cardiovascular outcomes in Table 1). Because providers submit a primary diagnosis code and up to nine secondary diagnosis codes and six procedure codes in each claim for a hospital stay, two approaches were considered: (1) calling an event only if a qualifying code was in the principal position or (2) calling an event if one of these same codes was in any position. Other factors, such as

variability in dates of diagnoses/procedures and Medicare identified events where WHI did not collect medical records, were explored.

Hlatky et al. examined agreement rates between fully adjudicated acute myocardial infarction and those identified by algorithms applied to Medicare data at the individual level.[7] The algorithm using only the principal diagnosis code provided excellent specificity (99.2%) but sensitivity and positive predictive value were not as strong (68% and 76% respectively, Table 2). Including acute myocardial infarction codes in any position improved sensitivity to 79% at a modest cost to specificity (98.8%) and positive predictive value (71%). In comparable person-level analyses, Lakshminarayan et al., found a similar pattern for stroke diagnoses with slightly higher sensitivity.[8] Mell et al. used an algorithm that accepted diagnosis codes in any position for abdominal aortic aneurysm, a much rarer condition, and found excellent specificity (greater than 99.9%) and very good sensitivity (91%) but the positive predictive value reached only 61%.[9] Kappa statistics ranged from 0.66 to 0.75.

The agreement rates for procedures were somewhat more robust than for diagnoses (Table 2), with sensitivity ranging from 76.4% to 94%, specificities all above 99.6% and positive predictive values ranging from 55.7% (lower extremity peripheral vascular procedure) to 90.3% (percutaneous coronary intervention).[7,9]

Discrepancies in identifying women with these clinical events between the two sources were examined to the extent possible in existing data. In over half of events found in Medicare but not in the WHI there were no corresponding self-reports of hospitalizations in WHI to trigger retrieval of medical records. Other reasons for these discrepancies included absence of medical records (no medical release obtained or protocol-defined reasons for not requesting documents) and confirmation of a related diagnosis. When an event was ascertained through WHI but no event was identified by the algorithm in Medicare, the underlying reasons for the discrepancies included: event of interest was not the principal diagnosis code, a related diagnosis code was used, an unrelated diagnosis code was used, or there were no claims for a hospitalization.[7-9] Dates of events from the two sources were very consistent (Table 3).[7-9]

Inference based on administrative data

The most important assessment of the value of administrative data for follow-up in pragmatic trials is the validity of the inference it is expected to support. The original WHI randomized trials provided an excellent venue for this evaluation because of the large number of participants enrolled in Medicare. Pooling the WHI hormone trials, Hlatky et al. found somewhat fewer clinical myocardial infarctions and non-procedure-related myocardial infarctions but more coronary revascularizations using Medicare data than were seen in the WHI data for these same participants. The resultant hazard ratios and 95% confidence intervals for the two data sources, however, were quite comparable (Table 4).[7]

Discussion

For selected cardiovascular events that involve hospitalization, the correspondence between events ascertained through Medicare and those identified through WHI protocol-driven mechanisms were reasonably well-aligned with those detected through the traditional methods of large-scale cohort studies. The specificity of the algorithms for these relatively low event rates conditions was excellent. Sensitivity and positive predictive value measures, while good, left room for improvement. For the purpose of assessing treatment effects in a randomized trial setting, the comparability of the hazard ratios and confidence intervals derived from administrative versus those from the original WHI randomized trial data provide reassurance that these methods can provide valid inference while reducing costs for pragmatic trials.

The similarity of the WHI hormone trial results is consistent with a growing literature from other cardiovascular trials suggesting that central adjudication does not add significant value to primary results of randomized trials when compared to investigator-reported,[11-13] registry-derived,[14] and death certificate[15] outcomes. Some have seen a critical impact of central adjudication on trial results,[16] however, and others have argued that adjudication is currently required by high impact journals.[17] Indeed, the controversial nature of the WHI hormone trial results suggested that compromises in data quality, regardless of its lack of effect on statistical measures, could have made these results even more contentious. A better appreciation of the robustness of randomized trial results may alleviate some of these concerns. Psaty et al.[18] found mostly quite consistent hazard ratios and confidence intervals for cardiovascular disease risk factors for myocardial infarction in the Cardiovascular Health Study when comparing Medicare derived and adjudicated outcomes, suggesting similar robustness in the context of a carefully-designed observational study. The smaller number of events detected through Medicare was a concern, however, especially if using only the principal diagnosis code. This would have implications for power that would need to be considered in a trial design. Floyd et al. reported a study of a rare adverse drug reaction in which there was a serious attenuation of the association between simvastatin and rhabdomyolysis with administrative data relative to validated cases in the context of an HMO.[19] The very low positive predictive power from their algorithm for administrative data is likely a strong factor in this poor performance.

In settings where outcomes misclassification is a concern, methods to improve inference have been proposed. For binary or discrete failure time data, adjustments incorporating estimates of sensitivity and specificity serve to de-attenuate the treatment effect estimates when misclassification is independent of treatment.[20-22] If misclassification depends on treatment or other covariates, however, bias in either direction could be introduced.[23] In this setting, one may improve treatment effect estimates and efficiency by incorporating a validation subsample within the larger study population to obtain and model the misclassification parameters.[20,22-26] In the context considered here, supplementation of administrative data with more traditional surveillance and outcomes assessments in a validation subsample would allow for estimation of performance statistics and adjustment of the treatment effect estimator.

More limited work has been done in the context of continuous failure time data. Dodd et al. propose an estimator that linearly combines treatment effect estimates from both sources in the validation subsample with that from the administrative data in the non-validation sample.[27] The estimates derived from administrative data are weighted by a factor determined to minimize the variance of the linear combination. The correlation between failure times in the administrative and the adjudicated data derived from the validation set is a key parameter in this weighting factor. Their estimator uses the adjudicated data to address bias in the administrative data and the administrative data to improve efficiency over a study using only adjudicated data. All of these proposed methods for analyzing outcomes with measurement error suggest that validation studies should be embedded in the trial design, at least until there is a better understanding of the statistical properties of claims data.

Some cautions regarding the use of administrative data are in order. Other outcomes than those examined here, particularly those lacking specific diagnosis codes (e.g., cancer recurrence) will be more challenging to assess through administrative data. For some cases, more complex algorithms may be needed. Chubak et al. examined multiple algorithms designed to optimize the performance measure of highest priority for assessing breast cancer recurrence, a condition without a diagnostic code.[28] Kroenke et al. extended this to use multiple algorithms (high sensitivity and high specificity) followed by medical record review for any discrepancies.[29] Such approaches may be useful for some outcomes with poorer documentation through insurance claims or when the performance characteristics are unacceptable.

The analyses here also point out the potential errors that may occur in traditional outcomes ascertainment in population-based studies. Although WHI was not able to trace back events detected only through Medicare, review of records and data already available to WHI suggest that there may be incomplete and/or incorrect recall of events, as well as administrative and logistical barriers that result in under-reporting of outcomes.

The close agreement of the hazard ratios and confidence intervals derived from adjudicated and administrative data here arose in the context of a randomized double-blinded trial with an active follow-up protocol. Randomization protects against imbalance on baseline factors but does not assure balance post-randomization. Differential outcomes ascertainment can occur when the trial is not blinded or when there are highly disparate treatment-related symptoms or side-effects (e.g., menopausal hormone increases vaginal bleeding, breast density and mammography performance[30] and finasteride effects on prostate-specific antigen levels and prostate size[31]). These treatment effects can lead to greater contact between participants and their health care providers or study staff or affect interpretation of test results, leading to differential surveillance, detection and/or diagnosis. Passive surveillance through administrative data sources may not protect against these biases. When these threats are present, a well-designed active follow-up protocol may be needed.

Practical matters in considering use of claims data include gaps in data availability, lag-time to receiving updates, and overall data complexity. The largest data loss from using Medicare linkage for WHI participants over 65 is in missing information for those who receive their care through HMOs. In younger US populations, reliance on claims data from a specific

insurer would be subject to similar gaps created by discontinuities in coverage. The WHISH trial is addressing this concern by design, requiring enrollment in fee-for-service Medicare as part of eligibility at baseline for the Self-Report Cohort and supplementing with targeted medical records review as needed to fill in gaps that occur over time. A more efficient approach for our research enterprise nationally would be to have access to HMO equivalent diagnostic and procedure information. Lack of detail for some conditions (e.g., procedure results, histologic classification) is another limitation of Medicare data.

The initial process of obtaining Medicare data requires advanced planning but once approved, the updates are more straightforward. These data, typically available annually, are less frequent and timely than usual for trial monitoring. In settings with significant safety concerns, reliance solely on administrative data to detect early adverse effects may be unacceptable. Medicare files are not easy to work with and the codes and format may change over time. The statistical analysis plan for a trial relying on Medicare data for outcomes needs flexibility to adapt to these changes that occur outside of the researchers' control.

In summary, pragmatic trials represent a critical step towards achieving evidence based medicine. Efforts to make such trials more affordable and feasible are needed so that more can be done within existing resources. Because claims data allow for much broader assessment of health outcomes and are available for a large population at significantly lower cost than any research driven data collection mechanisms, efforts to understand their value for inference for pragmatic trials are needed. In the WHI setting, the correspondence between selected cardiovascular events derived from Medicare claims to those ascertained through the WHI protocol was sufficiently promising to support the launching of WHISH, a new embedded trial that will rely heavily upon Medicare claims for outcomes data. The novel design features of WHISH, including a randomized consent design and a validation substudy may be a model for future pragmatic trials.

Acknowledgments

References

1. 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. [PubMed: 9492970]

2. Anderson GL, Manson J, Wallace R, et al. Implementation of the Women's Health Initiative study design. Ann Epidemiol. 2003; 13(9 Suppl):S5–S17. [PubMed: 14575938]

3. Curb JD, McTiernan A, Heckbert SR, et al. Outcomes ascertainment and adjudication methods in the Women's Health Initiative. Ann Epidemiol. 2003; 13(9 Suppl):S122–S128. [PubMed: 14575944]

4. ClinicalTrials.gov. [Accessed 31 May 2016] Women's Health Initiative Strong and Healthy Study (WHISH). https://clinicaltrials.gov/ct2/show/NCT02425345

5. National Institute on Aging at NIH. [Accessed 31 May 2016] Go4Life. https://go4life.nia.nih.gov/

6. Zelen M. A new design for randomized clinical trials. N Engl J Med. 1979; 300:1242–1245. [PubMed: 431682]

7. Hlatky MA, Ray RM, Burwen DR, et al. Use of Medicare data to identify coronary heart disease outcomes in the Women's Health Initiative. Circ Cardiovasc Qual Outcomes. 2014; 7:157–162. [PubMed: 24399330]

8. Lakshminarayan K, Larson JC, Virnig B, et al. Comparison of Medicare claims versus physician adjudication for identifying stroke outcomes in the Women's Health Initiative. Stroke. 2014; 45:815–821. [PubMed: 24525955]

9. Mell MW, Pettinger M, Proulx-Burns L, et al. Evaluation of Medicare claims data to ascertain peripheral vascular events in the Women's Health Initiative. J Vasc Surg. 2014; 60:98–105. [PubMed: 24636641]

10. Centers for Medicare & Medicaid Services. [Accessed 31 May 2016] Chronic Conditions Data Warehouse. https://www.ccwdata.org/web/guest/condition-categories

11. Granger CB, Vogel V, Cummings SR, et al. Do we need to adjudicate major clinical events? Clin Trials. 2008; 5:56–60. [PubMed: 18283081]

12. Pogue J, Walter SD, Yusuf S. Evaluating the benefit of event adjudication of cardiovascular outcomes in large simple RCTS. Clin Trials. 2009; 6:239–251. [PubMed: 19528133]

13. Hata J, Arima H, Zoungas S, et al. Effects of the endpoint adjudication process on the results of a randomized controlled trial: The ADVANCE trial. PLoS One. 2013; 8:e55807. [PubMed: 23390553]

14. Kjoller E, Hilde J, Winkel P, et al. Agreement between public register and adjudication committee outcome in a cardiovascular randomized clinical trial. Am Heart J. 2014; 169:197–204. 8. [PubMed: 25641528]

15. Doria-Rose VP, Marcus PM, Miller AB, et al. Does the source of death information affect cancer screening efficacy results? A study of the use of mortality review versus death certificates in four randomized trials. Clin Trials. 2010; 7:69–77. [PubMed: 20156958]

16. Serebruany VL, Atar D. Viewpoint: Central adjudication of myocardial infarction in outcome-driven clinical trials—Common patterns in TRITON, RECORD, and PLATO? Thromb Haemost. 2012; 108:412–414. [PubMed: 22836596]

17. Meinert L, Martin BK, McCaffrey LD, et al. Comment: Do we need to adjudicate major clinical events? Clin Trials. 2008; 5:557. [PubMed: 18827048]

18. Psaty BM, Delaney JA, Arnold AM, et al. Study of cardiovascular health outcomes in the era of claims data: The Cardiovascular Health Study. Circulation. 2015; 133:156–164. [PubMed: 26538580]

19. Floyd JS, Heckbert SR, Weiss NS, et al. Use of administrative data to estimate the incidence of statin-related rhabdomyolysis. JAMA. 2012; 307:1580–1582. [PubMed: 22511681]

20. Halloran ME, Longini IM Jr. Using validation sets for outcomes and exposure to infection in vaccine field studies. Am J Epidemiol. 2001; 154:391–398. [PubMed: 11532779]

21. Meier AS, Richardson BA, Hughes JP. Discrete proportional hazard models for mismeasured outcomes. Biometrics. 2003; 59:947–954. [PubMed: 14969473]

22. Magder LS, Hughes JP. Logistic regression when the outcome is measured with uncertainty. Am J Epidemiol. 1997; 146:195–203. [PubMed: 9230782]

23. Wang L, Shaw PA, Mathelier H, et al. Evaluating risk-prediction models using data from electronic health records. Ann Appl Stat. 2016; 10:286–304. [PubMed: 27158296]

24. Margaret AS. Incorporating validation subsets into discrete proportional hazards models for mismeasured outcomes. Stat Med. 2008; 27:5456–5470. [PubMed: 18613225]

25. Hunsberger S, Albert PS, Dodd L. Analysis of progression-free survival data using a discrete time survival model that incorporates measurements with and without diagnostic error. Clin Trials. 2010; 7:634–642. [PubMed: 21109582]

26. Edwards JK, Cole SR, Troester MA, et al. Accounting for misclassified outcomes in binary regression models using multiple imputation with internal validation data. Am J Epidem. 2013; 177:904–912.

27. Dodd LE, Korn EL, Freidlin B, et al. An audit strategy for progression-free survival. Biometrics. 2011; 67:1092–1099. [PubMed: 21210772]

28. Chubak J, Yu O, Pocobelli G, et al. Administrative data algorithms to identify second breast cancer events following early-stage Invasive breast cancer. J Natl Cancer Inst. 2012; 104:931–940. [PubMed: 22547340]

29. Kroenke CH, Chubak J, Johnson L, et al. Enhancing breast cancer recurrence algorithms through selective use of medical record data. J Natl Cancer Inst. 2016; 108:djv336. [PubMed: 26582243]

30. Chlebowski RT, Anderson G, Pettinger M, et al. Estrogen plus progestin and breast cancer detection by means of mammography and breast biopsy. Arch Intern Med. 2008; 168:370–377. [PubMed: 18299491]

31. Redman MW, Tangen CM, Goodman PJ, et al. Finasteride does not increase the risk of high-grade prostate cancer: a bias-adjusted modeling approach. Cancer Prev Res. 2008; 1:174–181.

Author Manuscript

Author Manuscript

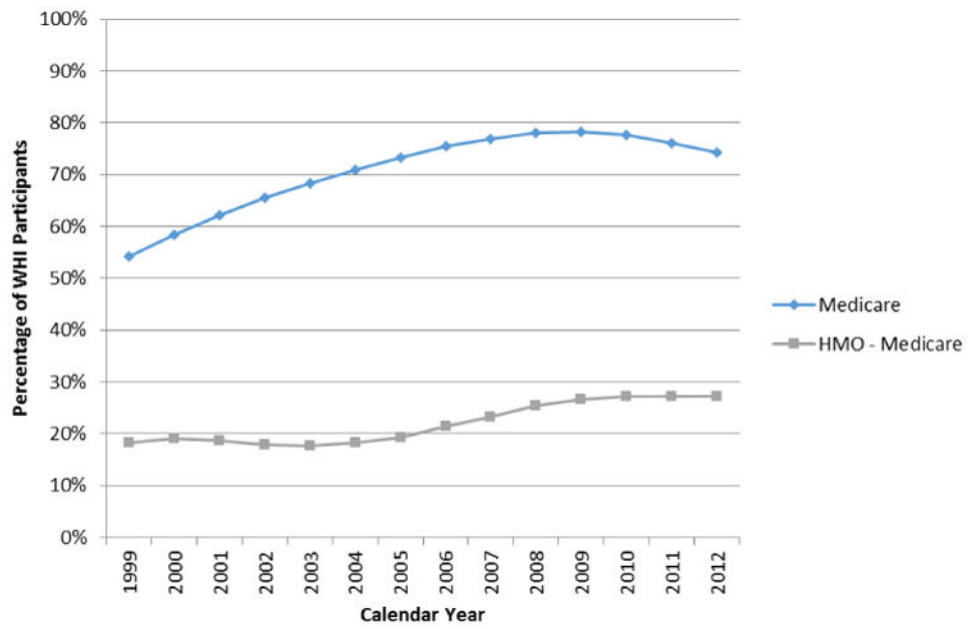Author Manuscript

Author Manuscript

Figure 1. Percentage of all WHI Participants (n=161,808) successfully identified as enrolled in Medicare overall and the subset of those receiving their Medicare coverage through a health maintenance organization (HMO), by calendar year

Table 1

Definitions of selected cardiovascular events in Medicare Inpatient (MedPar) files based on Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) Codes

Clinical event	Diagnostic and procedure codes used to define events in Medicare
Acute myocardial infarction	Hospital discharge codes: 410.×0, 410.×1
Coronary revascularization	Procedures codes for coronary artery bypass graft (36.1×, 36.2) or percutaneous coronary intervention (00.66, 36.0, 36.00, 36.01, 36.02, 36.05, 36.97)
Abdominal aortic aneurysm	Diagnosis codes (441.3-441.5, 441.9) or procedure codes (38.34,38.44,39.25,39.52,39.71) or CPT codes (35081, 35082, 35102, 35103, 35091, 0001T, 0002T, 35800, 34802-34805, 34830-34832)
Left Extremity peripheral artery disease	Diagnosis codes (440.20-440.24, 440.0, 443.0,444.22,444.81, 447.1,443.81 or 250.70), Procedure codes (39.50, 39.90, 00.55, 17.56, 99.10, 38.08,38.14, 38.16, 38.18, 39.25, 39.29, 84.11, 84.12, 84.15, 84.17) or CPT codes (35583,35556,35585,35571,35587,35621,35654,35646,35651,35656,35661, 35351,35355,35371,35302,35303,35306,34201,34203,35666,35665,35681, 27590,27880,27882,27884,28805,28810,28820,28825,37205,35471-35474,37226-37231,35470,35491,-35495,35606,37220,37221,37224,04 37225
Coronary artery Stenosis treatment	Diagnosis codes (433.10,433.30, 447.1, 433.11), procedure codes (38.42,38.12,39.90, 00.63) or CPT codes (35301,35390,37215,37216,0005T,0075T,0076T)

Table 2

Agreement between WHI adjudicated events and those defined by algorithms applied to CMS Medicare data

Outcome/CMS Algorithm	WHI-Yes CMS-Yes	WHI-No CMS-Yes	WHI-Yes CMS-No	WHI-No CMS-No	Sens %	Spec %	PPV %	NPV %	κ
Clinical Diagnoses									
Acute MI/Principal diagnosis code[7]	914	281	431	35771	68	99.2	76	98.8	0.71
Acute MI/Any diagnosis code[7]	1062	439	283	35613	79	98.8	71	99.2	0.74
Stroke/Principal diagnosis code[8]	458	182	124	23664	78.7	99.2	71.6	995	0.74
Stroke/Any diagnosis code[8]	505	240	77	23606	86.8	99.0	67.8	99.7	075
Abdominal aortic aneurysm[9]	32	20	13	49003	91.4	>99.9	61.5	>99.9	0.66
Clinical Procedures									
Coronary bypass surgery[7]	795	96	53	36453	94	99.7	89	99.9	0.91
Percutaneous coronary intervention[7]	1369	217	147	35664	86.3	99.6	90.3	99.4	0.88
Lower extremity peripheral vascular procedure[9]	175	139	54	50,253	76.4	99.7	55.7	99.9	0.64
Carotid artery treatment procedure[9]	320	77	27	48612	92.2	99.8	80.6	99.9	0.86

CMS Center for Medicare/Medicaid Services; Sens Sensitivity; Spec Specificity; PPV Positive Predictive Value; NPV Negative Predictive Value

Table 3

Correspondence between event dates derived from Medicare and those found through adjudicated medical records

Matching level	Acute MI[7]	Stroke[8]	Abdominal aortic aneurysm[9]	Lower extremity peripheral vascular procedures[9]	Carotid artery treatment[9]
Exact match	82%	83%	89%	75%	90%
Within ± 1 day		88%	89%	79%	93%
Within ± 7 days		95%	89%	82%	96%
Within ± 30 days	94%	96%	92%	85%	96%
Differs by exactly 365 days				0.7%	0.4%

Table 4

Hazard ratios and 95% confidence intervals from women age 65+ years and continuously enrolled in Medicare randomized to either active hormone therapy or placebo in either the WHI Estrogen plus Progestin or Estrogen-Alone trials

Clinical Event	Data source	Hormone Therapy (N=4142) N(%)	Placebo (N=4223) N(%)	Hazard Ratio	95% Confidence Interval
Acute myocardial infarction	WHI Adjudicated	150 (3.6)	120 (2.8)	1.31	(1.03-1.67)
	Medicare Algorithm	127 (3.1)	101 (2.4)	1.29	(1.00-1.68)
Non-procedure related myocardial infarction	WHI Adjudicated	130 (3.1)	105 (2.5)	1.28	(0.99-1.66)
	Medicare Algorithm	111 (2.7)	82 (1.9)	1.38	(1.03-1.83)
Coronary revascularization	WHI-Adjudicated	181 (4.4)	171 (4.1)	1.09	(0.88-1.35)
	Medicare Algorithm	195 (4.7)	179 (4.2)	1.10	(0.89-1.35)

Adapted from Hlatky et al.[7]