Published in final edited form as:

Med Care. 2014 May ; 52(5): e30–e38. doi:10.1097/MLR.0b013e31825a8c22.

Performance of Claims-Based Algorithms for Identifying Heart Failure and Cardiomyopathy Among Patients Diagnosed with Breast Cancer

Larry A. Allen, MD, MHS^{*}, Marianne Ulcickas Yood, DSC, MPH[†], Edward H. Wagner, MD, MPH[‡], Erin J. Aiello Bowles, MPH[‡], Roy Pardee, PhD[‡], Robert Wellman, MS[‡], Laurel Habel, PhD[§], Larissa Nekhlyudov, MD, MPH[¶], Robert L. Davis, MD, MPH^{||}, Onitilo Adedayo, MD, MSCR^{**}, and David J. Magid, MD, MPH^{††} for the Pharmacovigilance Research Group

^{*}Section of Advanced Heart Failure and Transplantation, Division of Cardiology, University of Colorado Anschutz Medical Center, Aurora, CO

[†]Department of Population Sciences, Henry Ford Hospital and Health System, Detroit, MI and Department of Epidemiology, Boston University School of Public Health, Boston, MA

[‡]Group Health Research Institute, Group Health Cooperative, Seattle, WA

Correspondence and Reprints: Larry A. Allen, MD, MHS, Academic Office 1, Room 7109, 12631 East 17th Avenue, Mailstop B130, PO Box 6511, Aurora, Colorado 80045. Phone: 303-724-4713. Fax: 303-724-2094. larry.allen@ucdenver.edu.

COMPLETE AUTHOR INFORMATION

1. **Larry A. Allen, MD, MHS**, Section of Advanced Heart Failure and Transplantation, Division of Cardiology, University of Colorado Anschutz Medical Center Academic Office 1, Room 7109, 12631 East 17th Avenue, Mailstop B130, PO Box 6511, Aurora, Colorado 80045. Phone: 303-724-4713. Fax: 303-724-2094. larry.allen@ucdenver.edu.
2. **Marianne Ulcickas Yood, DSC, MPH**, Department of Population Sciences, Henry Ford Hospital and Health System, Detroit, MI and Department of Epidemiology, Boston University School of Public Health, Boston, MA. muyood@muyood.com.
3. **Edward H. Wagner, MD, MPH**, Group Health Research Institute, Group Health Cooperative, 1730 Minor Ave, Suite 1600 Seattle WA 98101-1448. Phone: 206-287-2900. Fax: 206-287-2871. wagner.e@ghc.org.
4. **Erin J. Aiello Bowles, MPH**, Group Health Research Institute, Group Health Cooperative, Seattle, WA. 1730 Minor Ave, Suite 1600 Seattle WA 98101-1448. Phone: 206-287-2708. Fax: 206-287-2871. bowles.e@ghc.org.
5. **Roy Pardee, PhD**, Group Health Research Institute, Group Health Cooperative, 1730 Minor Ave, Suite 1600 Seattle WA 98101-1448. Phone: 206-287-2900. Fax: 206-287-2871. rpardee@gmail.com.
6. **Robert Wellman, MS**, Group Health Research Institute, Group Health Cooperative, 1730 Minor Ave, Suite 1600 Seattle WA 98101-1448. Phone: 206-287-2900. Fax: 206-287-2871. wellman.r@ghc.org.
7. **Laurel Habel, PhD**, Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612. Phone 510-891-3715. Fax: 510-891-3898. Laurel.Habel@kp.org.
8. **Larissa Nekhlyudov, MD, MPH**, Department of Population Medicine, Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, MA. larissa_nekhlyudov@harvardpilgrim.org.
9. **Robert L. Davis, MD, MPH**, Center for Health Research-Southeast, Kaiser Permanente Georgia, 11 Piedmont Center, 3495 Piedmont Road NE, Suite 110, Atlanta, GA 30305. Phone: 404-364-7197. Fax: 404-364-7361. Robert.L.Davis@nsmtp.kp.org.
10. **Onitilo Adedayo, MD, MSCR**, Marshfield Clinic Research Foundation, Marshfield Clinic, Marshfield, WI and Department of Hematology/Oncology, Weston, WI. onitilo.adedayo@marshfieldclinic.org.
11. **David J. Magid, MD, MPH**, Institute for Health Research, Kaiser Permanente Colorado, 10065 E. Harvard Ave Suite 300, Denver, CO 80231. Phone: 303.614.1300. djmagid@gmail.com.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

§Division of Research, Kaiser Permanente Northern California, Oakland, CA

¶Department of Population Medicine, Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, MA, and Department of Medicine, Harvard Vanguard Medical Associates, Boston, MA

||Center for Health Research-Southeast, Kaiser Permanente Georgia, Atlanta, GA

**Marshfield Clinic Research Foundation, Marshfield Clinic, Marshfield, WI and Department of Hematology/Oncology, Weston, WI

††Institute for Health Research, Kaiser Permanente Colorado, Denver, CO

Abstract

Background—Cardiotoxicity is a known complication of certain breast cancer therapies, but rates come from clinical trials with design features that limit external validity. The ability to accurately identify cardiotoxicity from administrative data would enhance safety information.

Objective—To characterize the performance of claims-based algorithms for identification of cardiac dysfunction in a cancer population.

Research Design—We sampled 400 charts among 6,460 women diagnosed with incident breast cancer, tumor size ≥ 2 cm or node positivity, treated within 8 US health care systems during 1999–2007. We abstracted medical records for clinical diagnoses of heart failure (HF) and cardiomyopathy (CM) or evidence of reduced left ventricular ejection fraction. We then assessed the performance of 3 different ICD-9-based algorithms.

Results—The HF/CM coding algorithm designed a priori to balance performance characteristics provided a sensitivity of 62% (95% confidence interval 40–80%), specificity of 99% (97–99%), positive predictive value (PPV) of 69% (45–85%), and negative predictive value (NPV) of 98% (96–99%). When applied only to incident HF/CM (ICD-9 codes and gold standard diagnosis both appearing after breast cancer diagnosis) in patients exposed to anthracycline and/or trastuzumab therapy the PPV was 42% (14–76%).

Conclusions—Claims-based algorithms have moderate sensitivity and high specificity for identifying HF/CM among patients with invasive breast cancer. Because the prevalence of HF/CM among the breast cancer population is low, ICD-9 codes have high NPV but only moderate PPV. These findings suggest a significant degree of misclassification due to HF/CM overcoding versus incomplete clinical documentation of HF/CM in the medical record.

Keywords

breast neoplasms; drug therapy; complications; cardiomyopathies; clinical coding

INTRODUCTION

Cardiotoxicity is a known complication of certain systemic chemotherapies.¹ Specifically, the development of cardiomyopathy (CM) with associated left ventricular systolic dysfunction (LVSD) and symptomatic heart failure (HF) are well documented adverse effects of anthracycline-based therapies and trastuzumab adjuvant therapy used to treat breast cancer.^{2–4} However, estimates of risk have been derived from randomized controlled trials (RCTs) with design features that may limit their external validity. Existing RCTs largely excluded older patients and those with comorbidities, especially pre-existing cardiovascular conditions that may potentially be associated with higher rates of drug toxicity. For example, of the 3 major trials of trastuzumab published in 2005 and 2006 that led to its widespread use, only 16% of women enrolled were 60 years or older^{5–7} compared

to more than 50% in the overall U.S. invasive breast cancer population and approximately 40% of the HER2-positive population.⁸ Additionally, all 3 of these trastuzumab trials excluded women with a history of HF, CM, coronary artery disease, uncontrolled hypertension, valvular disease, or unstable arrhythmias. More recent RCTs comparing the safety and efficacy of various chemotherapy regimens in HER-2 positive breast cancer have continued to enroll a majority of patients <50 years of age and exclude patients with major cardiovascular risk factors.⁹ Therefore, extrapolation of trial results to the routine care setting leaves major questions about real-world outcomes.¹⁰

Given the variability of treatment and greater vulnerability of patients with cancer treated in actual clinical practice, assessment of treatments and subsequent outcomes in community settings may improve our understanding of overall safety. The increasing use of health information technology potentially allows for automated identification and characterization of populations of patients with breast cancer, their treatments, and their subsequent outcomes from among large, contemporary, real-world datasets. Although diagnostic codes have been used to identify incident and prevalent HF among general patient cohorts,^{11–15} the accuracy of existing claims-based algorithms to identify HF/CM when applied to patients who also have a concurrent diagnosis of breast cancer is unknown. A variety of potential biases are possible. The major comorbidity of breast cancer has the potential to dominate claims-based coding leading to undercoding of HF/CM. Conversely, concerns about the possibility of chemotherapy-induced cardiotoxicity may result in overcoding of HF/CM. Previously published claims-based data analyses investigating cardiotoxicity among breast cancer patients have not validated their use of administrative codes,^{16–18} and consequently have drawn criticism.¹⁹

The multi-center Cancer Research Network (CRN)²⁰ offers an ideal setting for evaluating the performance of administrative diagnostic codes. The CRN is a consortium of 14 non-profit research centers based in integrated healthcare delivery organizations within the HMO Research Network.²⁰ Our objective was to test alternative strategies for identifying cardiotoxicity (i.e., symptomatic HF, CM, reduction in left ventricular ejection fraction [LVEF]) using electronic claims-based data.

METHODS

Patient Population

We included women diagnosed with incident invasive breast cancer from January 1, 1999 through December 31, 2007 who were enrolled in one of 8 CRN health plans: Group Health Cooperative, Harvard Pilgrim Health Care (restricted to women receiving care at Harvard Vanguard Medical Associates, a multispecialty integrated medical practice), Henry Ford Hospital and Health System, Marshfield Clinic, and Kaiser Permanent regions in Colorado, Georgia, Northern California, and Northwest. Women had to be enrolled for at least 12 months prior to breast cancer diagnosis, with censoring following death or disenrollment through December 31, 2008. Presence of invasive breast cancer was determined from existing tumor registries, as previously described.²⁰ The majority of data for this study were collected via the CRN Virtual Data Warehouse (VDW), which has been described in detail elsewhere.²⁰ This study was approved by the Group Health Institutional Review Board for Group Health and 5 other sites that ceded review to GH, and separately by the Institutional Review Boards at Marshfield Clinic and Henry Ford Health System.

To enrich the sample with patients more likely to receive chemotherapy, the eligible population (N=13,472) was further restricted to a final cohort women who had either tumor size ≥2cm or positive lymph nodes (N=6,460). We then used stratified random sampling to select women for detailed chart abstraction. The stratified random sampling algorithm

oversampled for patients 1) from small health systems, 2) with administrative codes for HF/CM, and 3) with exposure to certain chemotherapy (see Appendix 1 for sampling protocol). A total of 50 patients from each of the 8 CRN health care systems were chosen to create a chart abstraction cohort (N=400).

Administrative Codes for Heart Failure and Cardiomyopathy

Administrative codes chosen as potentially representing cardiotoxicity were all International Classification of Diseases, 9th Edition, Clinical Modification (ICD-9) diagnosis codes representing HF or CM: 398.91, 402.x1, 402.x3, 404.x1, 404.x3, 422.90, 425.4, 425.9, 428.xx. These codes were modeled on prior HF claims-based algorithms,^{12, 13} with the addition of the 425 “cardiomyopathy” codes due to the nature of cardiotoxicity, although some previously validated versions have included 425 coding.¹¹ Codes were available as part of the VDW, and those occurring 12 months before to 12 months after breast cancer diagnosis were included in this analysis. Although pharmacy data for common HF drugs and Physicians’ Current Procedural Terminology (CPT-4) codes HF-related imaging studies and procedures were initially considered for inclusion in the algorithm, preliminary analysis showed them to be overly non-specific (e.g. angiotensin converting enzyme inhibitors are primarily prescribed for hypertension) or highly insensitive (e.g. implantable cardioverter-defibrillator use was extremely rare). Furthermore, inclusion of LVEF measures into an automated algorithm was not feasible for the vast majority of cardiac imaging studies due to the free text nature of such data storage.

Chemotherapy Exposure

To ascertain detail on the type and timing of chemotherapy received, we used VDW procedure and pharmacy files (as well as non-VDW chemotherapy databases at some sites) to extract Healthcare Common Procedure Coding System and National Drug Codes specific to anthracyclines and trastuzumab with dates of administration. We also extracted CPT-4 codes related to chemotherapy infusion with dates of administration. All data were extracted up to 12 months after breast cancer diagnosis. Summary variables described whether women had any of the following treatment: anthracyclines, trastuzumab, other chemotherapy, unknown chemotherapy, or none.²¹

Manual Abstraction of the Medical Record

Trained abstractors reviewed the electronic and paper medical record from 12 months before breast cancer diagnosis to 12 months after breast cancer diagnosis. Abstractors received one initial 2-hour group training with a lead abstractor, and a subsequent 1-hour training after abstracting 10 records. Abstractors were instructed to review all chart notes within the 24-month window for any discussions of HF/CM (see Appendix 2 for a list of key words). Abstractors recorded verbatim all portions of the notes related to HF/CM, the date of visit, location of visit, and location within the chart. Abstractors were also instructed to record all cardiac imaging studies pertaining to LVEF measures within the 24-month window. Abstractors recorded the type of imaging (echocardiography, nuclear medicine, cardiac catheterization / ventriculogram, cardiac magnetic resonance, computerized tomography, or other), the date of each measurement, and the qualitative and quantitative systolic function measurements. Ten percent of charts from each CRN site were re-reviewed by the lead project manager for inconsistencies, inaccuracies, and missing data.

A practicing cardiologist board certified in advanced heart failure (Dr Allen) then reviewed all abstracted information to make a final assessment of the presence or absence of a clinical diagnosis of HF/CM. The Breast Cancer International Research Group definitions for cardiotoxicity,^{5–7,9} American College of Cardiology / American Heart Association standards for left ventricular dysfunction,²² and the European Society of Cardiology algorithm for HF

diagnosis²³ were used to provide standards for presence or absence of HF/CM. The final gold standard for HF/CM used in this analysis included either of the following: 1) a clinical diagnosis of HF/CM as determined from the medical record by cardiologist review; or 2) a documented quantitative LVEF of <50% or reduced qualitative LVEF (mildly, moderately, or severely). Patients were given a designation of “definite”, “indeterminate”, or “not” HF/CM. The “indeterminate” category was grouped with “not” HF/CM for the primary analyses, with secondary analyses considering these patients as HF/CM.

Prevalent versus Incident Disease

For the claims-based algorithms, patients with any HF/CM ICD-9 code present in the 12 months before the diagnosis of breast cancer were defined as having prevalent disease; patients with first HF/CM ICD-9 code identified in the 12 months after the diagnosis of breast cancer were classified as having incident HF/CM. Similar definitions were used for abstracted chart data, with the date of first mention of HF/CM or first LVEF <50% compared to the date of diagnosis of breast cancer. The date of breast cancer diagnosis was used instead of the date of exposure to anthracycline or trastuzumab therapy to create homogeneity between those receiving and those not receiving these therapies, as patients who do not receive therapy have no therapy start date from which to anchor post-therapy analyses. Because of possible detection biases after diagnosis of cancer (i.e. a diagnosis of cancer with consideration of cardiotoxic therapy may prompt a more thorough evaluation for pre-existing cardiac disease), we performed a secondary analysis whereby the cutoff between prevalent and incident disease was modified to date of exposure to anthracycline or trastuzumab therapy or 70 days after breast cancer diagnosis for unexposed patients (70 days was the median time to therapy initiation in exposed patients).

Claims-Based Algorithms

Three ICD-9 HF/CM code algorithms were created a priori: 1) one or more primary hospital discharge diagnoses of HF/CM, intended to provide a specific but potentially insensitive algorithm; 2) a HF/CM algorithm previously validated in the general population [1 primary hospital discharge diagnosis of HF/CM OR 3 secondary hospital discharge diagnoses of HF/CM OR 2 outpatient diagnoses of HF/CM (excluding ED) OR 3 ED HF/CM diagnoses OR 2 secondary hospital discharge + 1 outpatient diagnosis of HF/CM], intended to provide a balance between specificity and sensitivity;¹³ and 3) any ICD-9 code for HF/CM, intended to provide a sensitive but potentially non-specific algorithm.

Statistical Analysis

In order to assess the degree and direction of bias due to misclassification of HF/CM outcome, we estimated diagnostic accuracy measures for the claims-based algorithms. We compared the 3 ICD-9 code algorithms to a clinical gold standard derived from chart abstraction data. Due to our chart sampling scheme, the prevalence of HF/CM in the chart abstraction sample was higher than in the overall cohort. In order to standardize the results from the chart review sample back to the overall cohort, we calculated the inverse probability of verification given exposure, outcome, and LVEF status as indicated by the VDW.²⁴ Calculated weights were scaled to the sample size (n=400) to provide standard errors relative to the size of the validation cohort. Weighting was only applicable to women eligible for the chart review, which included 6460 women with positive lymph nodes and/or tumor size ≥2.0 cm. We calculated weighted sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) with 95% confidence intervals. Because of the importance of incident HF/CM in relationship to exposure to potentially cardiotoxic chemotherapy, we further assessed the diagnostic accuracy measures of the 3 different ICD-9 code algorithms for incident versus prevalent HF/CM and exposed versus non-exposed patients. All data analyses were performed using SAS (Cary, NC).

RESULTS

The initial cohort included 13,472 patients with incident, invasive breast cancer. The population eligible for chart review was limited to 6,460 women with either positive lymph nodes or tumors 2 cm or more in size, from which algorithm performance was determined. Among this final cohort, 36% of patients were >65 years of age, 54% received anthracyclines and/or trastuzumab, and 28% of patients received no chemotherapy. Compared to the chart-review eligible cohort, the 400 patients selected for detailed chart abstraction were older, had more comorbidity, had greater minority representation, had higher stage of breast cancer, and were more likely to receive chemotherapy than the original population (see Table 1).

Distribution of ICD-9 Codes for HF/CM Among Women with Breast Cancer

A primary hospital discharge diagnosis of HF/CM was found in 1.5% of the total cohort, and 10.8% of the oversampled chart abstraction cohort (Table 2). Outpatient ICD-9 codes for HF/CM occurred on 2 occasions in an additional 2.8% of the total cohort, and 18.0% of the oversampled chart abstraction cohort. Other combinations of ICD-9 codes constituting HF/CM in algorithm #2 were rare. Any ICD-9 code was found in 41.5% of the chart sample.

There was a more than 3-fold variation between sites in the frequency of ICD-9 codes for HF/CM among their breast cancer populations (data not shown). Among possible ICD-9 codes for HF/CM assigned to patients, the general code 428.0 “congestive heart failure, unspecified” was the most frequently used at each of the 8 sites. Codes that specified the acuity and type of heart failure (e.g. 428.21 “acute systolic heart failure”) were used 10-fold less frequently than the general 428.0 designation. Codes for cardiomyopathy, 425.4 and 425.9, accounted for 2.9–9.0% of HF/CM ICD-9 codes at individual sites. Hypertensive heart disease with HF (402 and 404) codes accounted for less than 1% of HF/CM ICD-9 codes.

Clinical and Imaging-Based Diagnosis of HF/CM Manually Abstracted from Charts

Description of HF/CM in the health record—Any mention of HF/CM (Appendix) was found in 122 of the 400 abstracted patients (30.5%), with a median of 3.5 unique episodes of care describing HF/CM per positively identified patient. On subsequent review by the study cardiologist, 92 were determined to be “definite” HF/CM, 6 “not” HF/CM, and 24 “indeterminate” due to inadequate documentation to apply strict HF/CM criteria (grouped as “not” for the primary analysis).

Cardiac imaging studies—Cardiac imaging studies with estimation of LVEF were found among 259 women (64.8%), with 73 patients having only 1 study, 51 patients with 2 studies, 36 having 3 studies, and 99 having 4–13 studies. Echocardiographic modalities were most common. Of the 400 sampled patients, 52 (13.0%) were found to have an LVEF recorded quantitatively at <50% or qualitatively as mildly, moderately, or severely reduced; 14 of 52 women (26.9%) had documentation of reduced LVEF without a clinical diagnosis of HF/CM in the chart.

Performance of Three ICD-9 Algorithms in the Breast Cancer Population

The sensitivity, specificity, PPV, and NPV are provided for the 3 different ICD-9-based claims-based algorithms in Table 3. When assessed against the combined gold standard of chart abstracted clinical diagnosis and/or reduced LVEF, the balanced algorithm (#2) had a sensitivity of 62% (95% confidence interval [CI] of 40–80%), specificity of 99% (97–99%), PPV of 69% (45–85%), and NPV of 98% (96–99%). Compared to the other algorithms, the balanced #2 algorithm maximized sensitivity and PPV with relatively similar specificity and

NPV; using only primary hospital discharge diagnosis ICD-9 codes for HF/CM resulted a sensitivity of only 14%, whereas using any ICD-9 HF/CM codes resulted in a PPV of only 45%. In secondary analyses, when we included patients with an “indeterminate” chart abstracted clinical diagnosis of HF/CM as disease true positive, the performance of algorithm #2 changed mildly: sensitivity 59% (40–76%), specificity 99% (97–100%), PPV 82% (58–94%), and NPV 97% (95–99%).

Algorithm Performance by Prevalent v. Incident and Exposed v. Unexposed

The sensitivity, specificity, PPV, and NPV for balanced algorithm #2 in relationship to prevalent and incident HF/CM are provided in Table 4. Mandating timing concordance between the appearances of claims-based and chart-abstracted HF/CM (i.e. both having first appearance before the date of breast cancer diagnosis [prevalent], or both having first appearance after the date of breast cancer diagnosis [incident]) had the expected effect of reduced sensitivity and PPV, though confidence intervals were wide due to decreased sample size. Specifically, of the 61 prevalent cases identified from medical record abstraction, 12 were classified as incident cases via the administrative claims algorithm.

Based on existing knowledge of the relationship of chemotherapy with cardiotoxicity, the ICD-9 algorithm performance characteristics of greatest research interest involve identification of incident HF/CM among patients exposed to anthracyclines and/or trastuzumab compared to the background rate of incident HF/CM. Algorithm performance for identification of incident HF/CM among exposed patients trended towards a higher PPV (Table 5); although due to smaller numbers of patients, the confidence intervals for performance characteristics are wide.

Secondary analysis showed that among the 29 exposed patients identified by the administrative claims algorithm as having incident HF/CM (Table 5), 5 patients had their ICD-9 code appear between cancer diagnosis and initiation of anthracycline / trastuzumab therapy; among the 27 unexposed patients identified by the algorithm as having incident HF/CM, 11 patients had their ICD-9 code appear within 70 days following cancer diagnosis. Similarly, among the 32 exposed patients identified by chart review as having incident HF/CM, 1 patient had diagnostic information appear between cancer diagnosis and initiation of anthracycline / trastuzumab therapy; among the 14 unexposed patients identified by chart abstraction as having incident HF/CM, 9 patients had their diagnostic information appear within 70 days following cancer diagnosis. Because of the even smaller number of incident cases identified under this delayed definition, algorithm performance characteristics for incident disease become unstable (95% CI for PPV 0–96.6%).

DISCUSSION

An algorithm of claims-based ICD-9 codes has moderate sensitivity and high specificity for HF/CM among women with incident invasive breast cancer. We found that the performance of an existing claims-based HF algorithm, modified here to include CM (ICD-9 codes 425.4 and 425.9), was lower in this unique population of breast cancer patients (overall PPV 69% [Table 3], incident-exposed PPV 42% [Table 5]) as compared to prior validation in a general population of adult patients (PPV 97%).^{11–13} These findings suggest a significant degree of misclassification for this automated HF/CM algorithm among women with invasive breast cancer, which reflects HF/CM overcoding (algorithm false positives) versus incomplete clinical documentation of HF/CM in the medical record (true disease missed by gold standard assessment).

Because the primary application of such an algorithm is assumed to be population-based estimates cardiac events associated with these drugs during real-world use, PPVs in the 40–

70% range give pause to such an approach. If overcoding is more common among patients after treatment with potentially cardiotoxic agents, significant overattribution from these observational associations is likely. However, because increased screening for cardiac disease is also likely to occur immediately after cancer diagnosis before initiation of chemotherapy, documentation of cardiac disease will then justify the avoidance of potentially cardiotoxic agents in such patients. Therefore, detection biases are likely to have complex implications for estimates of cardiotoxicity. Secondary analyses do suggest that new documentation and coding for HF/CM are particularly likely to occur soon after breast cancer diagnoses in patients who ultimately do not go on to receive anthracycline and/or trastuzumab therapy. Unfortunately, because of limited power in subgroup analyses it is impossible to provide precise estimates of the effect of these various biases.

Anthracycline and trastuzumab related cardiotoxicity is generally felt to manifest as cardiomyopathy documented as a decrease in the LVEF. Therefore, the cases of greatest interest among the broader cohort would be HF with new reductions in LVEF following exposure to these chemotherapies. Because of 1) the absence of serial cardiac imaging studies for the majority of patients, 2) the inability to automatically extract quantitative measures of LVEF from cardiac imaging studies within most electronic health records, and 3) the non-specific coding patterns for cardiac dysfunction (e.g. 428.0 “congestive heart failure, unspecified” accounted for the majority of HF/CM codes; codes such as 428.21 “acute systolic heart failure” were rare), we were forced to use an automated HF/CM algorithm that was indifferent to measures of left ventricular systolic function. Electronic databases in their present form and in the setting of current cardiac surveillance patterns are unlikely to allow for automated algorithms that can accurately distinguish between various types of cardiac dysfunction.

Algorithm decisions regarding trade-offs between sensitivity and specificity are often challenging and depend on the clinical setting to which the algorithm is applied. We can imagine situations where high specificity (reducing false positives) would be paramount and other situations where high sensitivity (reducing false negatives) would be the primary goal. Therefore, we investigated 3 different algorithms purposely designed to optimize different performance characteristics. We found that the balanced algorithm performed as expected (Table 3), providing comparatively good sensitivity and specificity, and therefore we emphasized it in the remainder of the subanalyses. However, future investigations may wish to use a highly sensitive or highly specific algorithm as circumstances dictate.

The variability in ICD-9 code frequencies for HF/CM within the 8 health delivery systems raises questions regarding the consistency of algorithm performance between institutions. There was nearly a 4-fold institutional difference in the frequency of both prevalent and incident HF/CM coding. However, these coding differences may, in part, reflect differences in institutional patient populations. Coding frequencies were concordant with the average age of the population treated by the institution (data not shown). For example, the site with the highest frequency of HF/CM codes treated a patient population of whom 51% were over 65 years of age, compared to the site with the lowest frequency of HF/CM codes which treated a patient population of whom only 22% were over 65 years of age. Different institutional culture for claims-based coding may provide an alternative explanation for at least part of these HF/CM coding differences. If the large degree of site-based variability in ICD-9 coding is not primarily a reflection of differences in case mix (e.g. age), this raises the concern that claims-based algorithms may have to be evaluated and interpreted independently for institutions with obviously different ICD-9 HF/CM coding practices.

Additional Limitations / Considerations

The manual review of all records by trained abstractors with subsequent review of only select abstracted data by a cardiologist is less rigorous than prior studies which have involved physicians for the entire process of chart abstraction. Other studies evaluating methods for identifying patients with HF have used a variety of standard definitions for HF, including Framingham criteria.²⁵ However, due to a lack of echocardiographic data on all patients and limitations in the naturalistic collection and recording of many of the criteria used for Framingham criteria (e.g., decrease in vital capacity by one third from maximum, nocturnal cough, S3 gallop, hepatojugular reflex), we chose to use provider assessments as recorded in the chart combined with what LVEF data was available to determine the presence or absence of HF/CM. In addition, the chart review was not prospectively designed to make a determination of prevalent versus incident HF/CM in relation to chemotherapy exposure (the distinction between prevalent and incident disease was before or after the date of breast cancer diagnosis, respectively). Finally, the collection of data in routine care lacks the standardization and detail that is inherent to the conduct of a clinical trial. Therefore, the gold standard presented here may have a greater degree of misclassification than seen in some prior studies, particularly for incident diagnoses. Prevalent disease was based on ascertainment within the 12 months prior to breast cancer diagnosis; it is conceivable that more prevalent disease would have been identified had we looked further backwards in time. We only reviewed claims and chart data in the 12 months after breast cancer diagnosis. Most adjuvant chemotherapy is given in the first year, and trastuzumab cardiotoxicity is thought to happen relatively acutely. However, it is well known that anthracycline cardiotoxicity can occur years after exposure. We did not assess for cardiotoxicity that developed more than a year after cancer diagnosis, and therefore coding algorithm performance is unknown for late-onset cardiotoxicity. Finally, due to the delayed and non-granular nature of administrative claims data, such an algorithm has limited utility in real-time identification of individual cardiac events.

Conclusion

The claims-based ICD-9 algorithms tested here had moderate PPVs for HF/CM. Therefore, claims-based algorithms in the setting of routine care, without the addition of some form of medical record review or other data enhancement, are crude tools for accurately estimating cardiotoxicity among women receiving chemotherapy for breast cancer. In the future, health care forces which promote greater standardization of ICD-9 coding and greater use of serial cardiac imaging with storage of LVEF measures in a way that allows for automated extraction from the electronic health record would improve the performance of automated electronic algorithms for characterizing cardiotoxicity of chemotherapy in community-based populations.

Acknowledgments

Funding: National Institutes of Health, National Cancer Institute supplement to the Cancer Research Network (U19 CA 79689)

REFERENCES

1. Felker GM, Adams KF Jr, Konstam MA, et al. The problem of decompensated heart failure: nomenclature, classification, and risk stratification. *Am Heart J.* 2003; 145(2 Suppl):S18–S25. [PubMed: 12594448]
2. Yeh ET, Bickford CL. Cardiovascular complications of cancer therapy: incidence, pathogenesis, diagnosis, and management. *J Am Coll Cardiol.* 2009; 53(24):2231–2247. [PubMed: 19520246]
3. Jones LW, Haykowsky MJ, Swartz JJ, et al. *J Am Coll Cardiol.* 2007; 50:1435–1441. [PubMed: 17919562]

4. Smith LA, Cornelius VR, Plummer CJ, et al. Cardiotoxicity of anthracycline agents for the treatment of cancer: systematic review and meta-analysis of randomised controlled trials. *BMC Cancer*. 2011; 10:337. [PubMed: 20587042]
5. Romond EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med*. 2005; 353:1673–1684. [PubMed: 16236738]
6. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med*. 2005; 353:1659–1672. [PubMed: 16236737]
7. Joensuu H, Kellokumpu-Lehtinen PL, Bono P, et al. Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. *N Engl J Med*. 2006; 354:809–820. [PubMed: 16495393]
8. Kurian AW, Fish K, Shema SJ, et al. Lifetime risks of specific breast cancer subtypes among women in four racial/ethnic groups. *Breast Cancer Res*. 2011; 12:R99. [PubMed: 21092082]
9. Slamon D, Eiermann W, Robert N, et al. Adjuvant trastuzumab in HER2-positive breast cancer. *N Engl J Med*. 2011; 365:1273–1283. [PubMed: 21991949]
10. Gross CP, Steiner CA, Bass EB, et al. Relation between prepublication release of clinical trial results and the practice of carotid endarterectomy. *JAMA*. 2000; 284:2886–2893. [PubMed: 11147985]
11. Ruo B, Capra AM, Jensvold NG, et al. Racial variation in the prevalence of atrial fibrillation among patients with heart failure: the Epidemiology, Practice, Outcomes, and Costs of Heart Failure (EPOCH) study. *J Am Coll Cardiol*. 2004; 43:429–435. [PubMed: 15013126]
12. Go AS, Yang J, Ackerson LM, et al. Hemoglobin level, chronic kidney disease, and the risks of death and hospitalization in adults with chronic heart failure: the Anemia in Chronic Heart Failure: Outcomes and Resource Utilization (ANCHOR) Study. *Circulation*. 2006; 113:2713–2723. [PubMed: 16754803]
13. Go AS, Lee WY, Yang J, et al. Statin therapy and risks for death and hospitalization in chronic heart failure. *JAMA*. 2006; 296:2105–2111. [PubMed: 17077375]
14. Lee DS, Donovan L, Austin PC, et al. Comparison of coding of heart failure and comorbidities in administrative and clinical data for use in outcomes research. *Med Care*. 2005; 43:182–188. [PubMed: 15655432]
15. Kumler T, Gislason GH, Kirk V, et al. Accuracy of a heart failure diagnosis in administrative registers. *Eur J Heart Fail*. 2008; 10:658–660. [PubMed: 18539522]
16. Doyle JJ, Neugut AI, Jacobson JS, et al. Chemotherapy and cardiotoxicity in older breast cancer patients: a population-based study. *J Clin Oncol*. 2005; 23:8597–8605. [PubMed: 16314622]
17. Pinder MC, Duan Z, Goodwin JS, et al. Congestive heart failure in older women treated with adjuvant anthracycline chemotherapy for breast cancer. *J Clin Oncol*. 2007; 25:3808–3815. [PubMed: 17664460]
18. Du XL, Xia R, Liu CC, et al. Cardiac toxicity associated with anthracycline-containing chemotherapy in older women with breast cancer. *Cancer*. 2009; 115:5296–5308. [PubMed: 19672997]
19. Portera CC, Swain SM. The heart of the matter. *J Clin Oncol*. 2007; 25:3794–3796. [PubMed: 17664459]
20. Wagner EH, Greene SM, Hart G, et al. Building a research consortium of large health systems: the Cancer Research Network. *J Natl Cancer Inst Monogr*. 2005; 3–11. [PubMed: 16287880]
21. Delate, T.; Aiello Bowles, E.; Pardee, T., et al. Validity of HMO administrative data for breast cancer chemotherapy exposure. Paper presented at, 17th Annual HMO Research Network Conference; March, 2011; Boston, MA.
22. Hendel RC, Budoff MJ, Cardella JF, et al. ACC/AHA/ACR/ASE/ASNC/HRS/NASCI/RSNA/SAIP/SCAI/SCCT/SCMR/SIR 2008 Key Data Elements and Definitions for Cardiac Imaging: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (Writing Committee to Develop Clinical Data Standards for Cardiac Imaging). *Circulation*. 2009; 119:154–186. [PubMed: 19064685]
23. Dickstein K, Cohen-Solal A, Filippatos G, et al. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2008: the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2008 of the European Society of Cardiology. Developed in collaboration

- with the Heart Failure Association of the ESC (HFA) and endorsed by the European Society of Intensive Care Medicine (ESICM). *Eur Heart J*. 2008; 29:2388–2442. [PubMed: 18799522]
24. Horvitz DG, Thompson DJ. A generalization of sampling without replacement from a finite universe. *J Am Stat Assoc*. 1952; 47:663.
25. McKee PA, Castelli WP, McNamara PM, et al. The natural history of congestive heart failure: the Framingham study. *N Engl J Med*. 1971; 285:1441–1446. [PubMed: 5122894]

Appendix 1

Stratification schema for sampling of patients for chart review at each of the 8 sites

Exposures	Outcomes						Patient Count
	Prevalent CM/HF		Incident CM/HF		No CM/HF		
	LVEF	No LVEF	LVEF	No LVEF	LVEF	No LVEF	
Both Anthracycline and Trastuzumab	1	1	2	2	3	1	10
Either Anthracycline or Trastuzumab	1	1	2	2	3	1	10
Neither Anthracycline nor Trastuzumab	3	3	6	6	9	3	30
Patient Count	5	5	10	10	15	5	50

CM – Cardiomyopathy, HF – Heart failure, LVEF – Left ventricular ejection fraction measured

Appendix 2

Words that chart abstractors were instructed to look for related to HF/CM

- Heart failure (HF)
- Congestive heart failure (CHF)
- Pulmonary congestion
- Pulmonary edema due to cardiac cause/etiology
- Left ventricular (LV) dysfunction
- Systolic dysfunction
- Diastolic dysfunction
- Low output/low cardiac output
- Cardiomyopathy (CM)
- Dilated cardiomyopathy
- Non-ischemic cardiomyopathy
- Ischemic cardiomyopathy
- Valvular cardiomyopathy
- Radiation-induced heart failure/myocarditis/pericarditis
- Chemotherapy (Adriamycin/anthracycline)-associated cardiomyopathy
- Herceptin (trastuzumab)-associated cardiomyopathy

Table 1

Characteristics of patients sampled for assessment of claims-based algorithm performance compared with eligible patients and full cohort

Characteristic	Chart-review patients (N=400)	Chart-review eligible patients (N=6,460)	All Patients (N=13,472)
Mean Age at Diagnosis (SD, median)	65 (14, 65)	60 (14, 58)	61 (13, 61)
Age ≥ 65 Years (n, %)	207 (52%)	2326 (36%)	5521 (41%)
Year of Cancer Diagnosis (n, %)			
1999	28 (7%)	644 (10%)	1281 (10%)
2000	40 (10%)	593 (9%)	1268 (9%)
2001	41 (10%)	720 (11%)	1595 (12%)
2002	44 (11%)	727 (11%)	1608 (12%)
2003	38 (10%)	714 (11%)	1520 (11%)
2004	42 (11%)	702 (11%)	1394 (10%)
2005	63 (16%)	769 (12%)	1564 (12%)
2006	59 (15%)	770 (12%)	1606 (12%)
2007	45 (11%)	821 (13%)	1636 (12%)
Charlson score			
0	229 (57%)	4506 (70%)	9340 (69%)
1	46 (12%)	883 (14%)	1925 (14%)
2	53 (13%)	640 (10%)	1393 (10%)
3+	72 (18%)	431 (7%)	814 (6%)
Race (n, %)			
White	299 (79%)	5331 (84%)	11312 (86%)
Black/African American	67 (18%)	735 (12%)	1336 (10%)
Other	15 (4%)	259 (4%)	550 (4%)
Worst AJCC stage (n, %)			
1	48 (16%)	520 (10%)	6354 (55%)
2a	126 (41%)	2575 (47%)	2601 (23%)
2b	56 (18%)	1108 (20%)	1110 (10%)
3/4	75 (25%)	1281 (23%)	1467 (13%)
Unknown	95	976	1940
Surgical treatment (n, %)	324 (81%)	6089 (95%)	12812 (95%)
Radiation treatment (n, %)	141 (39%)	3315 (42%)	7486 (57%)
Chemotherapy treatment			
None	157 (39%)	1827 (28%)	6240 (46%)
Anthracycline only	78 (20%)	2937 (46%)	3847 (27%)
Trastuzumab only	6 (2%)	98 (2%)	157 (1%)
Anthracycline + trastuzumab	74 (19%)	385 (6%)	470 (4%)
Other	85 (21%)	1213 (19%)	2748 (20%)

Table 2

Unweighted frequencies of heart failure and cardiomyopathy diagnoses by various claims-based and chart-review algorithms among the chart sample, eligible cohort, and full cohort

Outcome Algorithm	Chart-review patients (N=400)	Chart-review eligible patients (N=6,460)	All patients (N=13,472)
Virtual Data Warehouse (including claims-based data)			
Algorithm #1: 1 primary discharge diagnosis of HF/CM	43 (10.8%)	108 (1.7%)	199 (1.5%)
Algorithm #2:			
1 primary discharge OR	43 (10.8%)	108 (1.7%)	199 (1.5%)
3 secondary discharge OR	6 (1.5%)	20 (0.3%)	31 (0.2%)
2 outpatient OR	72 (18.0%)	206 (3.2%)	383 (2.8%)
3 ED OR	0	0	2 (0.0%)
2 secondary discharge + 1 outpatient diagnosis of HF/CM	3 (0.8%)	3 (0.1%)	8 (0.1%)
Any of the above	124 (31.0%)	337 (5.2%)	623 (4.6%)
Algorithm #3: Any ICD-9 code for HF/CM	166 (41.5%)	N/A	N/A
Medical Charts			
A) Clinician indication of HF/CM	92 (23.0%)	N/A	N/A
B) LVEF <50%	52 (13.0%)	N/A	N/A
C) Either A or B (gold standard)	108 (27.0%)	N/A	N/A

Table 3

Performance of ICD-9 coding algorithms for a diagnosis of heart failure or cardiomyopathy at any time during the study period when compared to chart-review

	AUTOMATED DATA HEART FAILURE EVENT DEFINITION		
	Algorithm #1 1 primary discharge dx of HF/CM N=43 (weighted N=5)	Algorithm #2 1 primary discharge OR 3 secondary discharge OR 2 outpatient OR 3 ED OR 2 secondary discharge + 1 outpatient diagnosis of HF/CM N=124 (weighted N=18)	Algorithm #3 Any ICD-9 code for HF/CM N=166 (weighted N=40)
(A) Clinical chart diagnosis of heart failure events N=92 (weighted N=17)	Sens = 16.6 (5.3–41.4) Spec = 99.5 (97.9–99.9) PPV = 59.3 (19.2–89.9) NPV = 96.4 (94.0–97.8)	Sens = 68.0 (43.6–85.4) Spec = 98.3 (96.3–99.2) PPV = 63.9 (40.6–82.1) NPV = 98.6 (96.7–99.4)	Sens = 93.0 (67.7–98.8) Spec = 93.7 (90.8–95.8) PPV = 40.0 (26.3–55.8) NPV = 99.7 (98.0–99.9)
(B) LVEF <50% N=52 (weighted N=8)	Sens = 20.6 (4.5–59.0) Spec = 99.2 (97.6–99.7) PPV = 34.1 (7.4–77.2) NPV = 98.4 (96.5–99.3)	Sens = 55.8 (23.8–83.5) Spec = 96.5 (94.1–97.9) PPV = 24.3 (10.0–48.2) NPV = 99.1 (97.4–99.7)	Sens = 79.6 (41.2–95.6) Spec = 91.4 (88.2–93.8) PPV = 15.9 (7.5–30.7) NPV = 99.6 (97.9–99.9)
(C) Either A or B (gold standard) N=108 (weighted N=21)	Sens = 14.0 (4.5–36.2) Spec = 99.5 (97.9–99.9) PPV = 59.3 (19.2–89.9) NPV = 95.5 (93.0–97.2)	Sens = 61.5 (39.6–79.5) Spec = 98.5 (96.6–99.3) PPV = 68.6 (44.9–85.4) NPV = 97.9 (95.9–99.0)	Sens = 87.7 (65.6–96.4) Spec = 94.2 (91.3–96.1) PPV = 44.9 (30.5–60.3) NPV = 99.3 (97.6–99.8)

Table 4

Performance of ICD-9 coding algorithms for a prevalent and incident heart failure or cardiomyopathy when compared to chart-review

ALL PATIENTS WITH BRCA	AUTOMATED ICD-9 ALGORITHM HF/CM EVENT	
	PREVALENT by Algorithm #2 (with first ICD-9 code occurring BEFORE date of breast cancer diagnosis) N=68 (weighted N=7)	INCIDENT by Algorithm #2 (with first ICD-9 code occurring AFTER date of breast cancer diagnosis) N=56 (weighted N=11)
PREVALENT by gold standard (with first chart mention of HF/CM or LVEF<50% occurring BEFORE date of breast cancer diagnosis) N=61 (weighted N=11)	Sens = 40.0 (16.9–68.9) Spec = 99.4 (97.8–99.8) PPV = 64.5 (27.8–89.6) NPV = 98.3 (96.4–99.2)	
INCIDENT by gold standard (with first chart mention of HF/CM or LVEF<50% occurring AFTER date of breast cancer diagnosis) N=46 (weighted N=9)		Sens = 41.7 (16.1–72.7) Spec = 98.1 (96.1–99.0) PPV = 33.3 (12.8–63.1) NPV = 98.6 (96.8–99.4)

Table 5

Performance of ICD-9 coding algorithms for incident heart failure or cardiomyopathy among those patients exposed to anthracycline and/or trastuzumab therapy

	ONLY EXPOSED PATIENTS	ONLY UNEXPOSED PATIENTS
	INCIDENT by Algorithm #2 (with first ICD-9 code occurring AFTER date of breast cancer diagnosis) N=29 (weighted N=7)	INCIDENT by Algorithm #2 (with first ICD-9 code occurring AFTER date of breast cancer diagnosis) N=27 (weighted N=4)
INCIDENT by gold standard (with first chart mention of HF/CM or LVEF<50% occurring BEFORE date of breast cancer diagnosis) N=32 exposed, 14 unexposed (weighted N= 5 exposed, 4 unexposed)	Sens = 57.8 (19.5–88.6) Spec = 98.1 (95.2–99.3) PPV = 42.1 (14.2–76.1) NPV =99.0 (96.3–99.7)	Sens = 19.9 (2.0–75.0) Spec = 98.0 (94.3–99.3) PPV = 18.3 (1.9–72.6) NPV = 98.2 (94.6–99.4)
INCIDENT by LVEF criteria (with first documented LVEF<50% occurring AFTER date of breast cancer diagnosis) N=28 exposed, 8 unexposed (weighted N= 4 exposed, 2 unexposed)	Sens = 55.3 (15.6–89.2) Spec = 97.8 (94.8–99.1) PPV = 33.0 (9.5–69.9) NPV =99.1 (96.5–99.8)	Sens = 4.7 (0.0–97.8) Spec = 97.6 (93.9–99.1) PPV = 2.1 (0.0–94.6) NPV = 99.0 (95.6–99.8)