JAMA Cardiology | Original Investigation

Accuracy of Medical Claims for Identifying Cardiovascular and Bleeding Events After Myocardial Infarction
A Secondary Analysis of the TRANSLATE-ACS Study

Patricia O. Guimarães, MD; Arun Krishnamoorthy, MD; Lisa A. Kaltenbach, MS; Kevin J. Anstrom, PhD; Mark B. Effron, MD; Daniel B. Mark, MD, MPH; Patrick L. McCollam, PharmD; Linda Davidson-Ray, MA; Eric D. Peterson, MD, MPH; Tracy Y. Wang, MD, MHS, MSc

IMPORTANCE Pragmatic clinical trial designs have proposed the use of medical claims data to ascertain clinical events; however, the accuracy of billed diagnoses in identifying potential events is unclear.

OBJECTIVES To compare the 1-year cumulative incidences of events when events were identified by medical claims vs by physician adjudication and to assess the accuracy of bill-identified events using physician adjudication as the criterion standard.

DESIGN, SETTING, AND PARTICIPANTS This post hoc analysis of a clinical trial assessed the medical claims forms and records for all rehospitalizations at 233 US hospitals within 1 year of the index acute myocardial infarction (MI) of 12 365 patients enrolled in the Treatment With Adenosine Diphosphate Receptor Inhibitors: Longitudinal Assessment of Treatment Patterns and Events After Acute Coronary Syndrome (TRANSLATE-ACS) study between April 1, 2010, and October 31, 2012. Fourteen patients (0.1%) died during the index hospitalization and were excluded from analysis. Recurrent MI, stroke, and bleeding events were identified per the *International Classification of Diseases, Ninth Revision, Clinical Modification* diagnosis and procedural codes in medical bills. These events were independently adjudicated by study physicians through medical record reviews using the prespecified criteria of recurrent MI and stroke and the bleeding definition by the Global Utilization of Streptokinase and Tissue Plasminogen Activator for Occluded Coronary Arteries (GUSTO) scale. Medical claims were reported on a Uniform Bill-04 claims form; claims were collected from all hospitals visited by patients enrolled in TRANSLATE-ACS. Agreement between medical claims–identified events and physician-adjudicated events over the 12 months after discharge was assessed with the κ statistic. Data were analyzed from January 30, 2015, to March 2, 2017.

MAIN OUTCOMES AND MEASURES Event rates within 1 year after MI.

RESULTS Among 12 365 patients with acute MI, 8890 (71.9%) were men and mean (SD) age was 60 (11.6) years. The cumulative 1-year incidence of events identified by medical claims was 4.3% for MI, 0.9% for stroke, and 5.0% for bleeding. Incidence rates based on physician adjudication were 4.7% for MI, 0.9% for stroke, and 5.4% for bleeding. Agreement between medical claims–identified and physician-adjudicated events was modest, with a κ of 0.76 (95% CI, 0.73 to 0.79) for MI and 0.55 (95% CI, 0.41 to 0.68) for stroke events. In contrast, agreement between medical claims–identified and physician-adjudicated bleeding events was poor, with a κ of 0.24 (95% CI, 0.19 to 0.30) for any hospitalized bleeding event and 0.15 (95% CI, 0.11 to 0.20) for moderate or severe bleeding on the GUSTO scale.

CONCLUSIONS AND RELEVANCE Event rates at 1 year after MI were lower for MI, stroke, and bleeding when medical claims were used to identify events than when adjudicated by physicians. Medical claims diagnoses were only modestly accurate in identifying MI and stroke admissions but had limited accuracy for bleeding events. An alternative approach may be needed to ensure good safety surveillance in cardiovascular studies.

TRIAL REGISTRATION clinicaltrials.gov Identifier: NCT01088503

JAMA Cardiol. 2017;2(7):750-757. doi:10.1001/jamacardio.2017.1460
Published online May 24, 2017.

Editor's Note page 758
Author Audio Interview
Supplemental content

Author Affiliations: Duke Clinical Research Institute, Duke University School of Medicine, Durham, North Carolina (Guimarães, Krishnamoorthy, Kaltenbach, Anstrom, Mark, Davidson-Ray, Peterson, Wang); John Ochsner Heart and Vascular Institute, Ochsner Medical Center, New Orleans, Louisiana (Effron); Eli Lilly & Company, Indianapolis, Indiana (McCollam).

Corresponding Author: Tracy Y. Wang, MD, MHS, MSc, Duke Clinical Research Institute, Duke University School of Medicine, 2400 Pratt St, Durham, NC 27705 (tracy.wang@duke.edu).

There is increasing interest in pragmatic clinical trials designed to compare therapeutic effectiveness and safety in routine clinical practice.[1-3] One proposed means of minimizing patient follow-up costs is to use medical claims as an alternative to traditional clinical follow-up and dedicated data collection to adjudicate events of interest.[4-7] Few studies have evaluated the accuracy of medical claims diagnosis codes when compared with the accuracy of physician event adjudication. Previous studies also have largely been limited to Medicare claims for patients older than 65 years.[8-10]

Treatment With Adenosine Diphosphate Receptor Inhibitors: Longitudinal Assessment of Treatment Patterns and Events After Acute Coronary Syndrome (TRANSLATE-ACS) was a multicenter, longitudinal study of 12 365 patients with acute myocardial infarction (MI) enrolled at 233 US hospitals.[11] Medical claims forms for all rehospitalizations of TRANSLATE-ACS participants during the study follow-up period (April 1, 2010, to May 13, 2014) were collected. Medical records were collected to perform independent physician adjudication of MI, stroke, and bleeding events. Our objectives were to (1) compare medical claims–identified vs physician-adjudicated cumulative incidence of recurrent MI, stroke, and bleeding events within 1 year after MI and (2) assess the accuracy of claims-identified events using physician adjudication as the criterion standard.

Key Points

Question Can medical claims be used to accurately assess cardiovascular and bleeding events after myocardial infarction in an all-ages population?

Findings In this secondary analysis of a clinical trial of 12 365 patients with acute myocardial infarction, the cumulative 1-year event rates for myocardial infarction, stroke, and bleeding were lower when medical claims were used to identify events compared with physician adjudication. Billed diagnoses were modestly accurate in identifying myocardial infarction and stroke admissions but had limited accuracy in identifying bleeding events.

Meaning The limited accuracy of medical claims in identifying bleeding events suggests the need for an alternative approach to ensure good safety surveillance in cardiovascular studies.

Methods

Patients and Study Design

The TRANSLATE-ACS was an observational study of 12 365 patients with acute MI treated with percutaneous coronary intervention and adenosine diphosphate receptor inhibitor therapy during the index MI hospitalization.[11] Patients were enrolled from April 1, 2010, through October 31, 2012. The institutional review board of the Duke University Health System approved this post hoc analysis. At enrollment, all patients provided written informed consent for medical claims and records collection related to any rehospitalization within the 15 months after index MI hospitalization. These claims and records were collected by trained personnel at the Duke Clinical Research Institute, Durham, North Carolina.[11,12]

Data Collection and Definitions

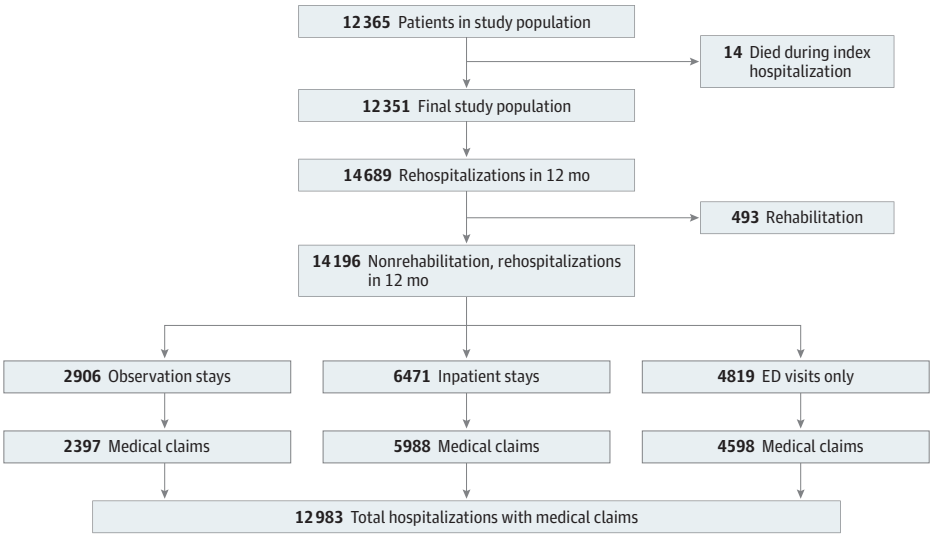
Data collection was triggered (1) by patient self-report of a rehospitalization, including screening of surrounding dates (within 7 days) and geographic areas (within 60 miles) if a hospital bill could not be found for the exact patient-reported date or location; (2) by additional hospital stays beyond the 7-day window reported by these hospitals; or (3) by standard queries of all TRANSLATE-ACS hospitals at 12 months after enrollment to screen for rehospitalizations that were underreported by the patient. Medical claims were reported on a Uniform Bill (UB-04) claims form, which is a common reporting format used by US hospitals that contains diagnosis and procedural codes. Hospital claims were collected from all hospitals visited by enrolled patients regardless of payer type. Medical claims reported only in-hospital deaths, whereas physician adjudication searched all TRANSLATE-ACS data resources to ascertain death, including death indexes, family contacts, site summary reports, and medical records.

Medical records, including discharge summaries, laboratory results, procedure reports, and operative reports, were collected after the bill confirmed at least an overnight rehospitalization. If the UB-04 bill could not be obtained or if *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*) codes were not available on the bill, then the study team requested records directly from sites to determine rehospitalization occurrence. The study physicians at the Duke Clinical Research Institute independently adjudicated outcomes after medical record review using the prespecified criteria of recurrent MI and stroke and bleeding definition by the Global Utilization of Streptokinase and Tissue Plasminogen Activator for Occluded Coronary Arteries (GUSTO) scale (eTable 1 in the Supplement).[13] For this analysis, the outcomes of recurrent MI, stroke, and bleeding were ascertained from medical claims using prespecified *ICD-9-CM* diagnosis and procedural codes (eTable 2 in the Supplement).[14-16] For each outcome, the primary analysis examined the first diagnosis code of each bill (primary diagnosis for the hospitalization). Secondary analyses were also conducted that considered diagnosis codes in the first or second position and then in any position on the bill. Notably, *ICD-9-CM* codes in medical claims identified bleeding events but could not classify GUSTO scale bleeding severity because *ICD-9-CM* codes do not capture data on hemoglobin drop or hemodynamic instability. Because transfusion can be identified by medical claims, secondary analyses examined the addition of transfusion to bleeding diagnosis codes.

Statistical Analysis

Data were analyzed from January 30, 2015, to March 2, 2017. We calculated the total number of each event type when identified by medical claims vs when physician adjudicated. We also calculated cumulative incidence rates at the patient level of each event type and the combined outcome of death, MI, and stroke when defined by the 2 respective methods. The primary analysis included the entire study population (N = 12 365). There were rehospitalizations for which medical claims could

Figure 1. Flowchart of Patients and Hospitalizations in the Treatment With Adenosine Diphosphate Receptor Inhibitors: Longitudinal Assessment of Treatment Patterns and Events After Acute Coronary Syndrome (TRANSLATE-ACS) Study



ED indicates emergency department.

not be collected or *ICD-9-CM* codes were unavailable on the bill. Thus, we performed sensitivity analyses to assess 1-year event rates after excluding patients with 2 or more rehospitalizations missing *ICD-9-CM* codes (n = 221) and after excluding patients with any rehospitalization missing *ICD-9-CM* codes (n = 762). We compared event rates using the first diagnosis code of medical claims and physician adjudication.

The agreement between medical claims-identified events and physician-adjudicated events over the 12 months after discharge was assessed with κ statistic, which compares the observed proportion of agreement between the 2 methods with the proportion of agreement expected by chance; a κ of 1 denotes complete agreement, and a κ of 0 demonstrates no more agreement than expected by chance. We repeated the κ statistic to examine agreement on a patient-level analysis that considered whether the patient had at least 1 of that event type within the year after discharge (eg, if a patient had 2 hospital stays with MI claims codes and only 1 stay was adjudicated as MI, then both claims and adjudication methods would agree at the patient level but not at the event level).

To examine potential discrepancies in analysis results when using claims-identified vs physician-adjudicated events, we compared 1-year rates of MI and bleeding as well as the composite of death, MI, and stroke between men and women. Myocardial infarction and bleeding cumulative incidence rates were compared between sexes using physician-adjudicated events data and claims data with codes in the first diagnosis position, first or second diagnosis position, and all diagnosis positions. We used Cox proportional hazards models to examine sex differences in event risk, presented as hazard ratios (HRs) and 95% CIs.

We used SAS version 9.4 (SAS Institute Inc) for all statistical analyses. $P < .05$ was considered statistically significant, and all tests were 2-tailed.

Results

Study Population and Hospitalizations

Among 12 365 patients, the mean (SD) age was 60 (11.6) years and 8890 (71.9%) were men. Of these, 14 patients (0.1%) died during the index hospitalization and were excluded (**Figure 1**). Among patients who were discharged alive (n = 12 351), 14 689 rehospitalizations occurred in the 12 months after index discharge; of these, 493 rehospitalizations (3.4%) were rehabilitation encounters and were excluded. Of the remaining rehospitalizations (n = 14 196), 2906 (20.5%) were observation stays involving at least 1 overnight stay, 6471 (45.6%) were inpatient stays, and 4819 (34.0%) were emergency department visits only. Bills with *ICD-9-CM* codes were available for 12 983 (91.5%) of the hospitalizations.

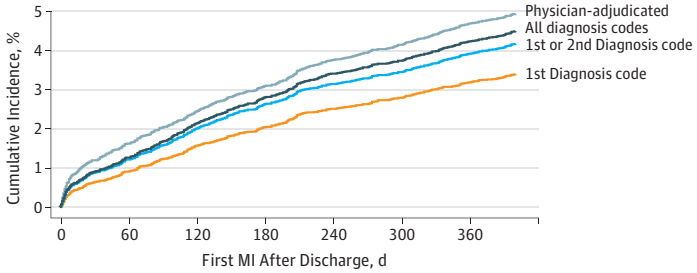
Incidence of Events

Cumulative incidence curves for physician-adjudicated and claims-identified clinical events are presented in **Figure 2**. The 1-year cumulative incidence rates for MI, stroke, and bleeding identified by the primary diagnosis code were generally lower with medical claims than with physician adjudication (**Table 1**). When broadened to all diagnosis codes, 1-year incidence rate of MI was still lower when ascertained by claims (4.3%) than by physician adjudication (4.7%), and so was the incidence of any bleeding (5.0% vs 5.4%). For stroke, both medical claims and physician adjudication rates were 0.9%. The rate of GUSTO moderate or severe bleeding by physician adjudication was 2.6%. The addition of transfusion codes to the bleeding definition overestimated the bleeding rates to 6.2%.

In sensitivity analyses excluding patients with 2 or more rehospitalizations missing *ICD-9-CM* codes (n = 221), the

Figure 2. Cumulative Incidence Curves

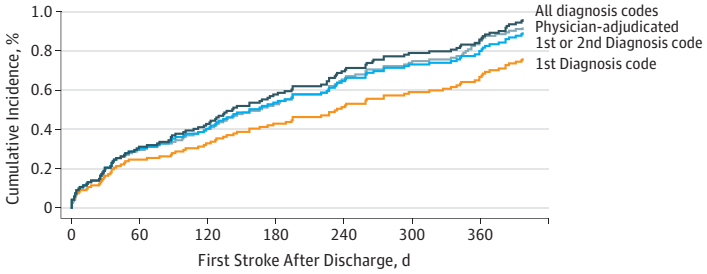
A Myocardial infarction



No. at risk

Physician-adjudicated	12351	11969	11783	11603	11446	11345	11179
1st Diagnosis code	12351	11975	11797	11621	11473	11372	11202
1st or 2nd Diagnosis code	12351	12010	11846	11681	11536	11436	11272
All diagnosis codes	12351	11935	11759	11581	11417	11309	11140

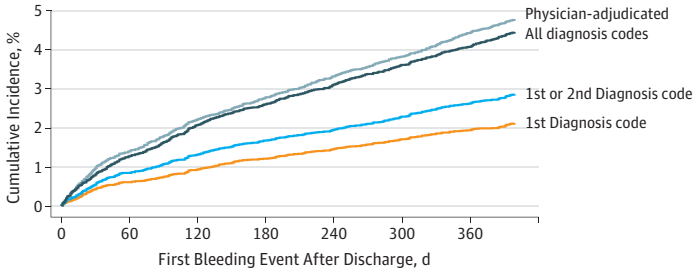
B Stroke



No. at risk

Physician-adjudicated	12351	12085	11985	11856	11748	11668	11536
1st Diagnosis code	12351	12086	11987	11859	11751	11671	11539
1st or 2nd Diagnosis code	12351	12090	11992	11868	11868	11682	11549
All diagnosis codes	12351	12087	11989	11863	11863	11675	11539

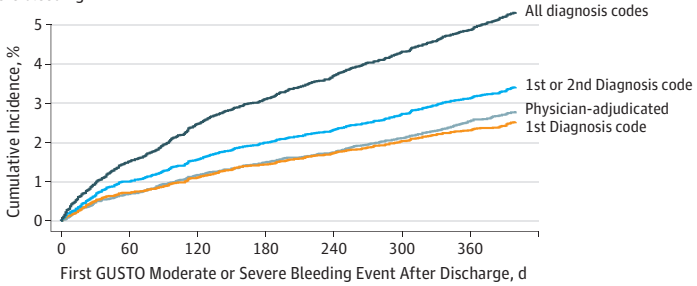
C Any bleeding



No. at risk

Physician-adjudicated	12351	11944	11749	11572	11409	11277	11097
1st Diagnosis code	12351	12001	11850	11692	11560	11447	11282
1st or 2nd Diagnosis code	12351	12033	11900	11752	11624	11520	11368
All diagnosis codes	12351	11923	11732	11555	11384	11251	11059

D GUSTO moderate or severe bleeding



No. at risk

Physician-adjudicated	12351	11944	11749	11572	11409	11277	11097
1st Diagnosis code	12351	12001	11850	11692	11560	11447	11282
1st or 2nd Diagnosis code	12351	12033	11900	11752	11624	11520	11368
All diagnosis codes	12351	12041	11904	11763	11637	11531	11366

Cumulative incidence curves for physician-adjudicated and medical claims–identified myocardial infarction (MI) (A), stroke (B), any bleeding event (C), and Global Utilization of Streptokinase and Tissue Plasminogen Activator for Occluded Coronary Arteries (GUSTO) scale moderate or severe bleeding event (D).

cumulative incidence rates of physician-adjudicated events were similar to the primary analysis in all study patients; however, the incidence of claims-identified events was lower, particularly for bleeding (eTable 3 in the Supplement). When patients with any rehospitalizations missing *ICD-9-CM* codes (n = 762) were excluded, similar results were found.

Hospital bills were only able to ascertain in-hospital death, identifying 195 of 431 total deaths (45.2%) confirmed among patients discharged alive after the index MI. The incidence of the composite outcome of death, MI, and stroke at 1 year was 7.5% by physician adjudication; by medical claims, this rate was 5.1% when only the primary diagnosis code was used and 6.0% when all diagnosis codes were used.

Table 1. Cumulative Incidence of Medical Claims–Identified and Physician-Adjudicated Events Within 1 Year

Events	Incidence Rate, %	
	Medical Claims	Physician Adjudication
Myocardial infarction		
First diagnosis code	3.2	4.7
First or second diagnosis code	4.0	4.7
All diagnosis codes	4.3	4.7
Stroke		
First diagnosis code	0.7	0.9
First or second diagnosis code	0.8	0.9
All diagnosis codes	0.9	0.9
All hospitalized bleeding events		
First diagnosis code	2.4	5.4
First or second diagnosis code	3.2	5.4
All diagnosis codes	5.0	5.4
All bleeding with inclusion of transfusion codes		
First diagnosis code	4.4	5.4
First or second diagnosis code	5.0	5.4
All diagnosis codes	6.2	5.4
Death/myocardial infarction/stroke		
First diagnosis code	5.1	7.5
First or second diagnosis code	5.8	7.5
All diagnosis codes	6.0	7.5

Event-Level Agreement

Table 2 displays agreement between claims-identified events and physician-adjudicated events. For MI, agreement between medical claims and physician adjudication was reasonable when only the first diagnosis code of each bill was examined (κ = 0.62; 95% CI, 0.59-0.66) and improved when the examination was broadened to include all diagnosis codes in the bill (κ = 0.76; 95% CI, 0.73-0.79). Agreement was lower for stroke (κ = 0.52; 95% CI, 0.39-0.64) and improved marginally by using the first or second diagnosis codes (κ = 0.59; 95% CI, 0.46-0.72) or all diagnosis codes (κ = 0.55; 95% CI, 0.41-0.68).

Agreement for bleeding events was markedly lower; the κ was 0.22 (95% CI, 0.18-0.25) and did not improve substantially when the search was expanded to the first or second diagnosis codes (κ = 0.25; 95% CI, 0.21-0.30) or to all diagnosis codes (κ = 0.24; 95% CI, 0.19-0.30). When compared with physician-adjudicated GUSTO moderate or severe bleeding, the

Table 2. Agreement Between Medical Claims–Identified and Physician-Adjudicated Events

Event	No. of Events by Type of Agreement				κ (95% CI)[a]
	Claims Yes, Physician Yes	Claims No, Physician Yes	Claims Yes, Physician No	Claims No, Physician No	
Myocardial infarction					
First diagnosis code	482	264	66	1145	0.62 (0.59 to 0.66)
First or second diagnosis code	588	158	90	1121	0.73 (0.70 to 0.76)
All diagnosis codes	625	121	103	1108	0.76 (0.73 to 0.79)
Stroke					
First diagnosis code	101	28	12	42	0.52 (0.39 to 0.64)
First or second diagnosis code	115	14	17	37	0.59 (0.46 to 0.72)
All diagnosis codes	120	9	23	31	0.55 (0.41 to 0.68)
All bleeding					
First diagnosis code	351	514	31	302	0.22 (0.18 to 0.25)
First or second diagnosis code	431	434	52	281	0.25 (0.21 to 0.30)
All diagnosis codes	575	290	129	204	0.24 (0.19 to 0.30)
All bleeding, including transfusion codes					
First diagnosis code	565	300	211	122	0.18 (−0.04 to 0.07)
First or second diagnosis code	612	253	228	105	0.02 (−0.04 to 0.08)
All diagnosis codes	690	175	273	60	−0.02 (−0.08 to 0.03)
GUSTO scale moderate or severe bleeding					
First diagnosis code	150	241	232	575	0.10 (0.04 to 0.15)
First or second diagnosis code	198	193	285	522	0.14 (0.09 to 0.20)
All diagnosis codes	279	112	425	382	0.15 (0.11 to 0.20)

Abbreviation: GUSTO, Global Utilization of Streptokinase and Tissue Plasminogen Activator for Occluded Coronary Arteries.

[a] All *P* values <.001.

agreement of medical claims worsened in both the primary analysis (κ = 0.10; 95% CI, 0.04-0.15) and when diagnosis codes in all positions were considered (κ = 0.15; 95% CI, 0.11-0.20). Agreement was even worse when adding transfusion codes to the definition of bleeding (κ = 0.18; 95% CI, −0.04 to 0.07).

Patient-Level Agreement

Patient-level agreement was better than event-level agreement (**Table 3**). Patient-level agreement improved when the search from diagnosis codes in the first position was expanded to all diagnosis codes. For MI, the κ was 0.71 (95% CI, 0.68-0.74) when only the primary diagnosis code was considered, 0.79 (95% CI, 0.76-0.82) for the code in the first or second position, and 0.82 (95% CI, 0.79-0.84) for all diagnosis codes. The κ reached 0.87 (95% CI, 0.82-0.92) for stroke and 0.68 (95% CI, 0.65-0.71) for all bleeding when all diagnosis codes were considered. When transfusion codes were added to the definition of bleeding, agreement between medical claims and physician adjudication was better; the κ was 0.68 (95% CI, 0.65-0.71) for the primary diagnosis code and 0.70 (95% CI, 0.67-0.73) for all diagnosis codes. For the composite outcome of death, MI, and stroke, the κ was 0.72 (95% CI, 0.70-0.75) for the diagnosis code in the first position, 0.77 (95% CI, 0.75-0.79) for the code in the first or second position, and 0.78 (95% CI, 0.76-0.80) for the code in all diagnosis positions.

To assess the differences in analysis results using claims vs trials data with physician-adjudicated events, we compared the risks of MI and bleeding as well as the composite of death, MI, and stroke at 1 year between women and men. For all outcomes, event rates were higher among women; the absolute difference in event rates between women and men were slightly larger when ascertained by physician adjudication than by medical claims (eTable 4 in the Supplement). Regardless of event ascertainment method, women had a higher cumulative incidence of MI than men (6.0% vs 4.2% by physician adjudication; 5.3% vs 3.9% by medical claims using all diagnosis codes). The HRs comparing the MI risk of women with that of men were similar (HR, 1.43 [95% CI, 1.22-1.67] by physician adjudication; HR, 1.37 [95% CI, 1.16-1.64] by medical claims using all diagnosis codes). The HRs were closest when medical claims using all diagnosis code positions were compared with physician adjudication. Similar results were observed for bleeding (7.2% vs 4.7% by physician adjudication; 6.6% vs 4.3% by medical claims using all diagnosis codes) and for the composite of death, MI, and stroke (9.2% vs 6.9% by physician adjudication; 7.5% vs 5.4% by medical claims using all diagnosis codes).

Table 3. Correlation Between Medical Claims–Identified and Physician-Adjudicated Events in Event-Level and Patient-Level Analyses

Events	κ (95% CI)	
	Event Level	Patient Level
Myocardial infarction		
First diagnosis code	0.62 (0.59 to 0.66)	0.71 (0.68 to 0.74)
First or second diagnosis code	0.73 (0.70 to 0.76)	0.79 (0.76 to 0.82)
All diagnosis codes	0.76 (0.73 to 0.79)	0.82 (0.79 to 0.84)
Stroke		
First diagnosis code	0.52 (0.39 to 0.64)	0.83 (0.77 to 0.89)
First or second diagnosis code	0.59 (0.46 to 0.72)	0.87 (0.82 to 0.92)
All diagnosis codes	0.55 (0.41 to 0.68)	0.87 (0.82 to 0.92)
All bleeding		
First diagnosis code	0.22 (0.18 to 0.25)	0.56 (0.52 to 0.60)
First or second diagnosis code	0.25 (0.21 to 0.30)	0.62 (0.59 to 0.66)
All diagnosis codes	0.24 (0.19 to 0.30)	0.68 (0.65 to 0.71)
All bleeding, including transfusion codes		
First diagnosis code	0.18 (−0.04 to 0.07)	0.68 (0.65 to 0.71)
First or second diagnosis code	0.02 (−0.04 to 0.08)	0.70 (0.67 to 0.73)
All diagnosis codes	−0.02 (−0.08 to 0.03)	0.70 (0.67 to 0.73)
GUSTO scale moderate or severe bleeding		
First diagnosis code	0.10 (0.04 to 0.15)	0.43 (0.38 to 0.48)
First or second diagnosis code	0.14 (0.08 to 0.20)	0.46 (0.42 to 0.51)
All diagnosis codes	0.15 (0.11 to 0.20)	0.48 (0.44 to 0.52)

Abbreviation: GUSTO, Global Utilization of Streptokinase and Tissue Plasminogen Activator for Occluded Coronary Arteries.

Discussion

Pragmatic clinical trial designs have proposed using preexisting data, such as medical claims, to ascertain clinical events as a cost-efficient alternative to dedicated trial-specific follow-up. Our study provides insight into the value and limitations of using billing data in clinical studies. We found (1) lower event rates when identified by medical claims rather than by physician adjudication, (2) moderate event-level agreement for MI and stroke when ascertained by medical claims rather than by physician adjudication of medical records but poor agreement for bleeding events, and (3) much better patient-level agreement than event-level agreement for MI, stroke, and bleeding outcomes. Our results recommend exercising caution when using medical claims data alone for event identification, particularly for bleeding events.

Estimated event rates based on billing data were lower than those based on adjudication; therefore, some events may be missed when using medical claims as the sole method of outcomes assessment. Claims data may miss events for several reasons. First, medical bills may not always be collectable. In TRANSLATE-ACS, claims containing diagnosis code information were collected for 12 983 rehospitalizations (91.5%), with lower rates of observation-visit collection, even if the rehospitalization involved an overnight stay, than the rates for inpatient admissions. Second, medical claims can detect only billable events. Coding practices are driven by reimbursement and not focused on comprehensive event documentation. Third, medical claims can ascertain only in-hospital events. Serious events often entail hospitalization, but early mortality may be missed; only 195 deaths (45.2%) in TRANSLATE-ACS were in-hospital events identified by medical bill data. Finally, even a

comprehensive claims data source (eg, insurance payer data) or electronic health record will miss events if the patient transitions in and out of health plans or health systems or if the platform does not actively track vital status.

Claims data have been used increasingly to assess clinical outcomes in comparative effectiveness studies.[17] A previous study linked data from the Women's Health Initiative to Medicare claims for patients older than 65 years and found that MI event agreement between medical claims and physician adjudication was reasonable, with a κ of 0.71 when only the first diagnosis code was examined and a κ of 0.74 when the first 2 diagnosis codes were examined.[9] A similar analysis was performed for stroke,[10] which showed moderate agreement between Medicare data and adjudicated events (κ = 0.69). Moreover, several studies have shown a high positive predictive value of using Medicare claims data to capture ischemic events.[8,9,15,18] Medicare is a unique data source because Medicare beneficiaries seldom disenroll. Consequently, ascertainment of vital status and billed clinical events is fairly comprehensive. Nevertheless, Medicare data can describe only patients older than 65 years enrolled in fee-for-service programs, limiting the data's use in clinical trials that also enroll patients younger than 65 years or patients with or without private health insurance.

The TRANSLATE-ACS study offered a unique opportunity to evaluate claims data accuracy because its protocol mandated bill collection for all patients regardless of age and payer status. Furthermore, all outcome events were centrally adjudicated. We found moderate event-level agreement between primary diagnoses of medical claims and physician adjudication for MI and stroke events. Because coding priorities for medical claims are based on reimbursement potential, the hospitalization primary diagnosis code may not align with events of interest. When we expanded the search to all diagnosis codes, the agreement rates improved.

Cardiovascular trials often need to identify bleeding events for safety surveillance. Previous studies have observed high positive predictive values for identifying bleeding events with administrative data, especially intracranial and gastrointestinal tract hemorrhages, using medical record abstraction as the criterion standard.[19,20] In this study, we observed poor event-level agreement between medical claims and physician adjudication for bleeding events even when all diagnosis codes were considered (κ = 0.24). Although bleeding severity cannot be ascertained from billed codes, severe bleeding would more likely be a billable diagnosis, but the κ was only 0.15 for GUSTO moderate or severe bleeding when all diagnosis codes were considered. Patient-level agreement was better than event-level agreement but was still lower than the agreement observed for MI and stroke. Moreover, agreement did not substantially improve when transfusion codes were added to the definition of bleeding. For example, we compared differences in MI and bleeding risk between women and men using 2 data sources. Women consistently had higher cumulative incidence of events than did men, but the direction of the sex effect in MI and bleeding risk, as well as the composite of death, MI, and stroke, was the same despite the method of event ascertainment. The HR point estimates were closest between medical claims and physician adjudication when all diagnosis code positions in the claims data set were considered.

Patient-level analyses had better correlation than event-level analyses, which is reassuring because patient-level analyses are commonly conducted in clinical trials. In our study, patient-level κ values were 0.80 or higher for recurrent MI and stroke, suggesting that medical bill data may be a reasonable approach to ascertain ischemic outcomes in clinical studies. Nevertheless, given lower event-level correlation, time-to-event analyses may have differing results when medical claims are used rather than physician adjudication. Our findings suggest exercising caution when using medical claims to assess bleeding events because patient-level correlation is lower than that for MI or stroke, and event-level correlations are poor. On the basis of these results, we recommend alternative approaches beyond medical claims to validate bleeding events and to appropriately identify bleeding of greater severity that may be more clinically relevant.

Limitations

The results of this study should be interpreted in light of some limitations. First, some medical claims (n = 37) could not be collected from hospitals despite multiple attempts because of patient privacy concerns (despite patient-signed informed consent and medical record release forms) and the lack of resources to pull information for research purposes. UB-04 forms containing diagnosis code data were not available from certain hospital types and encounters, particularly observation-status hospitalizations. Therefore, several sensitivity analyses excluded patients for whom *ICD-9-CM* code information could not be obtained. Second, medical bill codes may have been inaccurate. For example, manual review of select cases that disagreed revealed that early readmissions after the index MI hospitalization had a diagnosis code of 410.x1 (acute MI, initial encounter), but medical record review revealed that this code alluded to the previous hospitalization event. Furthermore, the process of physician adjudication can be imperfect. Finally, because of the lack of a US universal health record, we cannot exclude the possibility that hospitalizations may have been missed despite the extensive screening of rehospitalizations and safeguards built into the study design to ascertain events.

Conclusions

Event rates at 1 year were lower for MI, stroke, and bleeding when medical claims were used rather than physician adjudication. Moderate agreement between medical claims and physician adjudication was observed in ascertaining MI and stroke events, but agreement was worse for bleeding events. While medical claims may be a reasonable resource to assess MI and stroke outcomes, caution is still needed. Medical claims have limited accuracy in identifying bleeding events, which suggests the need for an alternative approach to ensure good safety surveillance in cardiovascular studies.

ARTICLE INFORMATION

Accepted for Publication: April 3, 2017.

Published Online: May 24, 2017. doi:10.1001/jamacardio.2017.1460

Author Contributions: Drs Guimarães and Wang had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analyses.
Study concept and design: Guimarães, Krishnamoorthy, Davidson-Ray, Wang.
Acquisition, analysis, or interpretation of data: Guimarães, Krishnamoorthy, Kaltenbach, Anstrom, Effron, Mark, McCollam, Peterson, Wang.
Drafting of the manuscript: Guimarães, McCollam, Wang.
Critical revision of the manuscript for important intellectual content: Krishnamoorthy, Kaltenbach, Anstrom, Effron, Mark, McCollam, Davidson-Ray, Peterson, Wang.
Statistical analysis: Guimarães, Kaltenbach, Anstrom, Peterson.
Obtained funding: Mark, McCollam, Wang.
Administrative, technical, or material support: Effron, McCollam, Davidson-Ray, Wang.
Study supervision: McCollam, Peterson, Wang.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Krishnamoorthy reported working on projects for the Duke Clinical Research Institute (DCRI) funded by research grants from the National Institutes of Health, Novartis, Daiichi Sankyo, and Eli Lilly & Company and attended educational conferences by HeartWare, Thoratec, and Medtronic. Dr Anstrom reported receiving significant research grant support from AstraZeneca and Eli Lilly & Company and modest fees for consulting and other non–continuing medical education services from Janssen Pharmaceutical. Dr Effron reported being a current shareholder and former employee of Eli Lilly & Company. Dr Mark reported having consulted for Medtronic, CardioDx, and St Jude Medical and received research grants from Eli Lilly & Company, Gilead, AstraZeneca, AGA Medical, Bristol-Myers Squibb, Merck & Co, and Oxygen Therapeutics. Dr McCollam reported being a current shareholder and a full-time employee of Eli Lilly & Company. Dr Peterson reported receiving grant support from American College of Cardiology, American Heart Association, and Janssen as well as consulting fees from Bayer, Boehringer Ingelheim, Merck, Valeant, Sanofi, AstraZeneca, Janssen, Regeneron Pharmaceuticals, and Genentech. Dr Wang reported working on projects for the DCRI funded by AstraZeneca, Boston Scientific, Bristol-Myers Squibb, Daiichi Sankyo, Eli Lilly & Company, Gilead Sciences, GlaxoSmithKline, and Regeneron and receiving consulting honoraria from AstraZeneca, Eli Lilly & Company, Merck, and Pfizer. No other disclosures were reported.

Funding/Support: This study was funded in part by grant U01HL107023 from the National Heart, Lung, and Blood Institute and by fellowship funding from Conselho Nacional de Desenvolvimento Cientifico e Tecnologico, Ministry of Science, Technology and Innovation of Brazil (Dr Guimarães). TRANSLATE-ACS is sponsored by Daiichi Sankyo and Lilly USA. The Duke Clinical Research Institute is the coordinating center for this study, representing a collaborative effort with the American College of Cardiology.

Role of the Funder/Sponsor: The funding sources reviewed the manuscript prior to submission, but submission decisions were made by the coauthors.

Additional Contributions: Erin Campbell, MS, Duke University, provided editorial assistance and prepared the manuscript. She did not receive compensation for her contribution.

REFERENCES

1. Califf RM, Woodlief LH. Pragmatic and mechanistic trials. *Eur Heart J*. 1997;18(3):367-370.

2. Sugarman J, Califf RM. Ethics and regulatory complexities for pragmatic clinical trials. *JAMA*. 2014;311(23):2381-2382.

3. Sedgwick P. Explanatory trials versus pragmatic trials. *BMJ*. 2014;349:g6694.

4. Mahaffey KW, Harrington RA, Akkerhuis M, et al; For the PURSUIT Investigators. Systematic adjudication of myocardial infarction end-points in an international clinical trial. *Curr Control Trials Cardiovasc Med*. 2001;2(4):180-186.

5. Mahaffey KW, Wampole JL, Stebbins A, et al; APEX-AMI Investigators. Strategic lessons from the clinical event classification process for the Assessment of Pexelizumab in Acute Myocardial Infarction (APEX-AMI) trial. *Contemp Clin Trials*. 2011;32(2):178-187.

6. Pogue J, Walter SD, Yusuf S. Evaluating the benefit of event adjudication of cardiovascular outcomes in large simple RCTs. *Clin Trials*. 2009;6(3):239-251.

7. Granger CB, Vogel V, Cummings SR, et al. Do we need to adjudicate major clinical events? *Clin Trials*. 2008;5(1):56-60.

8. Psaty BM, Delaney JA, Arnold AM, et al. Study of cardiovascular health outcomes in the era of claims data: the Cardiovascular Health Study. *Circulation*. 2016;133(2):156-164.

9. 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(1):157-162.

10. 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(3):815-821.

11. Chin CT, Wang TY, Anstrom KJ, et al. Treatment with adenosine diphosphate receptor inhibitors-longitudinal assessment of treatment patterns and events after acute coronary syndrome (TRANSLATE-ACS) study design: expanding the paradigm of longitudinal observational research. *Am Heart J*. 2011;162(5):844-851.

12. Krishnamoorthy A, Peterson ED, Knight JD, et al. How reliable are patient-reported rehospitalizations? implications for the design of future practical clinical studies. *J Am Heart Assoc*. 2016;5(1):e002695. doi:10.1161/JAHA.115.002695

13. GUSTO Investigators. An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. *N Engl J Med*. 1993;329(10):673-682.

14. Fosbol EL, Wang TY, Li S, et al. Warfarin use among older atrial fibrillation patients with non–ST-segment elevation myocardial infarction managed with coronary stenting and dual antiplatelet therapy. *Am Heart J*. 2013;166(5):864-870.

15. Kiyota Y, Schneeweiss S, Glynn RJ, Cannuscio CC, Avorn J, Solomon DH. Accuracy of Medicare claims–based diagnosis of acute myocardial infarction: estimating positive predictive value on the basis of review of hospital records. *Am Heart J*. 2004;148(1):99-104.

16. Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF. Accuracy of *ICD-9-CM* codes for identifying cardiovascular and stroke risk factors. *Med Care*. 2005;43(5):480-485.

17. Mentz RJ, Hernandez AF, Berdan LG, et al. Good clinical practice guidance and pragmatic clinical trials: balancing the best of both worlds. *Circulation*. 2016;133(9):872-880.

18. Kumamaru H, Judd SE, Curtis JR, et al. Validity of claims-based stroke algorithms in contemporary Medicare data: Reasons for Geographic and Racial Differences in Stroke (REGARDS) study linked with Medicare claims. *Circ Cardiovasc Qual Outcomes*. 2014;7(4):611-619.

19. Al-Ani F, Shariff S, Siqueira L, Seyam A, Lazo-Langner A. Identifying venous thromboembolism and major bleeding in emergency room discharges using administrative data. *Thromb Res*. 2015;136(6):1195-1198.

20. Arnason T, Wells PS, van Walraven C, Forster AJ. Accuracy of coding for possible warfarin complications in hospital discharge abstracts. *Thromb Res*. 2006;118(2):253-262.