ORIGINAL RESEARCH ARTICLE

Pharmacoeconomics 2011; 29 (8): 693-704
1170-7690/11/0008-0693/$49.95/0

Cost Estimation of Cardiovascular Disease Events in the US

Amy K. O'Sullivan,[1] *Jaime Rubin,*[1] *Joshua Nyambose,*[1] *Andreas Kuznik,*[2] *David J. Cohen*[3] and *David Thompson*[1]

1 i3 Innovus, Medford, Massachusetts, USA
2 Pfizer Inc, New York, New York, USA
3 Cardiovascular Research, St Luke's Mid-America Heart Institute, Kansas City, Missouri, USA

Abstract

Background: In this study, we developed cost prediction equations that facilitate estimation of the costs of various cardiovascular events for patients of specific demographic and clinical characteristics over varying time horizons.
Methods: We used administrative claims data and generalized linear models to develop cost prediction equations for selected cardiovascular events, including myocardial infarction (MI), angina, strokes and revascularization procedures. Separate equations were estimated for patients with events and for their propensity score-matched controls. Attributable costs were estimated on a monthly basis for the first 36 months after each event and annually thereafter, with differences in survival between cases and controls factored into the longitudinal cost calculations. The regression models were used to estimate event costs ($US, year 2007 values) for the 'average' patient in each event group, over various time periods ranging from 1 month to lifetime.
Results: When the equations are run for the average patient in each event group, attributable costs of each event in the acute phase (i.e. first 3 years) are substantial (e.g. MI $US73 300; hospitalization for angina $US36 000; nonfatal haemorrhagic stroke $US71 600). Furthermore, for most events, cumulative costs remain substantially higher among cases than among controls over the remaining lifetime of the patients.
Conclusions: This study provides updated estimates of medical care costs of cardiovascular events among a managed care population over various time horizons. Results suggest that the economic burden of cardiovascular disease is substantial, both in the acute phase as well as over the longer term.

Background

Cardiovascular disease (CVD) has been well studied in the US since the initiation of the Framingham Heart Study over 60 years ago. One of the most well known and longest running epidemiological studies in the world, the Framingham Heart Study recently enrolled its third generation of participants.[1] The study has been instrumental in identifying risk factors for CVD and

establishing the roles of hypertension, dyslipidaemia, diabetes, smoking and other risk factors in the development of the disease.[1] Data from the study have been used to estimate risk prediction equations for various CVD events for patients with specific demographic and clinical characteristics.[2-4]

Other predictive models in CVD have been developed over the years, including the Coronary Heart Disease Policy Model in the US,[5] the Cardiovascular Life Expectancy Model in Canada[6] and the Coronary Heart Disease (CHD) Policy Analysis in England and Wales.[7,8] Economic studies in CVD have ranged from incidence-based cost-of-illness studies[9-12] to economic evaluations of primary and secondary prevention interventions.[13-16] The latter include economic evaluations conducted alongside cardiovascular endpoint trials[13,16] and cost-effectiveness modelling studies.[15,16] Both types of analyses require incidence-based cost estimates of the clinical endpoints of interest (e.g. myocardial infarction [MI], stroke) for estimation of economic impact of the intervention under study. Researchers obtain these costs from the published literature or, in some cases, an analysis of healthcare claims data.

While mean cost estimates obtained from the literature are useful for economic evaluation, these become quickly outdated; moreover, it is also the case that the costs of CVD events can vary substantially according to patient demographic and clinical characteristics. In this study, we developed a series of cost prediction equations that facilitate estimation of the costs of various cardiovascular events for patients of specific demographic and clinical characteristics over varying time horizons. This paper describes the methods and data used to construct the equations and presents the results for our patient population. A technical appendix containing the beta-coefficients from the equations is available in the Supplemental Digital Content 1, http://links.adisonline.com/PCZ/A119.

Methods

Overview

We used administrative claims data from a large US health plan to estimate econometric equations that predict medical care costs of 15 different cardiovascular events based on available demographic and clinical information. The equations also include time-since-event-onset variables, which permit estimation of costs incurred over varying time periods of interest, ranging from month of onset to lifetime. We estimated separate equations for each of the following CVD events: MI (fatal and non-fatal), cardiac arrest (fatal and resuscitated), congestive heart failure (CHF), angina pectoris, transient ischaemic attack (TIA), haemorrhagic stroke (fatal and non-fatal), ischaemic stroke (fatal and non-fatal), peripheral vascular disease (PVD), coronary artery bypass graft (CABG) surgery and percutaneous transluminal coronary angioplasty (PTCA) separately with and without stent. For each event except TIA, patients were required to be hospitalized for the event. Hospitalization was confirmed by ensuring that the relevant primary diagnosis code was reported on either a facility or a physician claim occurring during a hospital stay.

The study used standard techniques of incidence-based cost-of-illness estimation.[10,17] Conceptually, such techniques involve longitudinal estimation of costs attributable to a health event from a base year of onset until disease resolution or through all remaining future years of life (whichever comes first), incorporating all relevant medical care services received along the way. These include not only the acute phase treatment for the health event and intermediate care received during recovery and follow-up, but also treatment of secondary events and other disease conditions to which the primary event may give rise.

Due to patterns of disenrolment from the plan, there was insufficient follow-up for most patients to conduct a lifetime cost analysis using a single cohort. Therefore, we used different methods to estimate medical care costs during the first 3 years following event onset ('acute-phase analysis') versus thereafter ('post-year 3 analysis'). In the acute-phase analysis, an econometric, cross-sectional time-series approach was employed, while the post-year 3 analysis was cross-sectional only and involved estimation of an average annual cost over the patient's remaining lifetime.

Data Sources

We estimated the cost prediction equations using medical and pharmacy claims data for commercial and managed Medicare enrolees from a large US health plan affiliated with i3 Innovus, an Ingenix Inc. company. The study population was drawn from a database containing information on approximately 21.5 million commercial and Medicare Advantage members from 1 January 2002 through 31 December 2006. Membership in this health plan is geographically diverse across the US, and beneficiaries have fully insured coverage for physician, hospital and prescription drug services. Providers of services submit their payment claims directly to the health plan. For patients with more than one insurer, claims are typically observed by both health plans and handled through a coordination of benefits process. This assures that most or all of the claims for these patients are observed in the database. Claims data are de-identified, and each enrolee is assigned a randomly generated unique identifier before being placed in the database. The study was conducted in accordance with established corporate guidelines for adherence to Health Insurance Portability and Accountability Act (HIPAA) privacy requirements, and review by an institutional review board was not sought.

For purposes of estimating post-event survival probabilities, data on relative risks (RRs) of death among those with versus those without events of interest were obtained from three community-based cohort studies.[18-20] These studies provide data on the RR of death for ischaemic stroke (RR = 2.3)[18] and hazard ratios (HRs) for death due to CHF (HR = 2.1),[20] angina (diagnosed, HR = 3.0; symptomatic, HR = 2.0)[19] and MI (HR = 3.7).[19] In the absence of comparable data for haemorrhagic stroke, we assumed its HR was the same as that for ischaemic stroke based on clinical input; similarly, the HR for death for resuscitated cardiac arrest was assumed to be the same as that for CHF, and the HRs for TIA, PVD, CABG and PTCA were assumed to be equal to that of the 'symptomatic angina' group from Mosterd et al.[20] RRs and HRs were used to calculate excess hazards of death for each event; excess hazards were then applied to age- and sex-specific background HRs for death from US life tables to arrive at survival probabilities for those with CVD events. Based on clinical input, excess risks of death associated with CVD events were assumed to last 10 years post-event, after which patients were assumed to revert back to age- and sex-specific general population mortality risks. Post-event survival probabilities were used in the post-year 3 analysis only, as fatalities were included in the acute-phase sample.

Sample Selection

Patients were included in the acute-phase analysis if they were aged ≥35 years in calendar year 2003 and had a claim for a study event any time during that year. Cardiovascular events were identified in the claims data using *International Classification of Diseases*, *Ninth Revision*, *Clinical Modification* (ICD-9-CM) codes, ICD-9-CM procedure codes, Current Procedural Terminology (CPT) codes and Health Care Financing Agency Common Procedure Coding System (HCPCS) codes. The specific codes used to identify patients are presented in table I.

An event was defined as fatal if the primary diagnosis field on either a facility or physician claim occurring during a hospital stay contained a relevant code, the hospital discharge status was coded as 'dead', and there was no evidence of healthcare claims occurring >30 days after discharge (to ensure no miscoding of discharge status). Conversely, non-fatal events required evidence that the patient was alive at discharge (i.e. hospital discharge status not equal to dead and healthcare claims occurring after discharge).

For each person in the database, the earliest observed event (chronologically) was considered his or her index study event; no analyses of subsequent events were performed. These persons made up the candidate 'cases' for each event; the date of the event was defined as the patient's 'index date'. In instances in which multiple events occur simultaneously (e.g. hospitalization for MI in which revascularization takes place), the event was included as a case for the procedure (e.g. revascularization) rather than the diagnosis (MI).

Table I. Identification of cardiovascular disease (CVD) events

CVD event	ICD-9-CM	ICD-9 procedure	CPT-4	HCPCS
Fatal and non-fatal MI	410.xx			
Fatal and resuscitated cardiac arrest	427.5			
Hospitalization for CHF	398.91, 402.01, 402.11, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, 428.xx			
Hospitalization for (worsening) angina	411.1, 411.8x, 413.x			
TIA	362.34, 435.x			
Fatal and non-fatal haemorrhagic stroke	430, 431, 432.9			
Fatal and non-fatal ischaemic stroke	433.01, 433.11, 433.21, 433.31, 433.81, 433.91, 434.01, 434.11, 434.91, 436			
Hospitalization for PVD[a]	405.01, 405.11, 405.91, 440.xx–442.xx, 443.1–443.2x, 443.9, 444.x–445.x, 447.1	00.03, 00.55, 00.61, 00.63–00.64, 38.02–38.06, 38.08, 38.12–38.16, 38.18, 38.32–38.36, 38.38, 38.42–38.46, 38.48, 39.22–39.26, 39.28–39.29, 39.50–39.52, 39.7x, 39.90	34001–203, 34800–35152, 35301–90, 35450–9, 35470–5, 35480–95, 35501–87, 35601–71, 35691–5, 35700, 35875–907, 37184–6, 37201, 37205–8, 37215–6, 93668	
CABG surgery[b]		36.10–36.14	33510–23, 33533–6	S2205–9
PTCA with stent[c]		00.66, 36.01–36.02, 36.05 and 36.06 or 36.07	92982–4 or 92995–6 or 92980–1	
PTCA without stent[d]		00.66, 36.01–36.02, 36.05		

a Patients were identified as a case for hospitalization for PVD if they had relevant ICD-9-CM *or* relevant ICD-9 Procedure codes *or* relevant CPT-4 codes.

b Patients were identified as a case for CABG if they had relevant ICD-9 Procedure codes *or* relevant CPT-4 codes *or* relevant HCPCS codes.

c Patients were identified as a case for PTCA with stent if they had relevant ICD-9 Procedure codes for PTCA *or* relevant CPT-4 codes for physician services rendered *and* either a procedure code for stent *or* a code to insert a stent.

d Patients were identified as a case for PTCA without stent if they had a relevant ICD-9 Procedure code for PTCA *or* relevant CPT-4 code for physician services rendered.

CABG = coronary artery bypass graft; **CHF** = congestive heart failure; **CPT** = Current Procedural Terminology; **HCPCS** = Health Care Financing Agency Common Procedure Coding System; **ICD-9-CM** = *International Classification of Diseases, Ninth Revision, Clinical Modification*; **MI** = myocardial infarction; **PTCA** = percutaneous transluminal coronary angioplasty; **PVD** = peripheral vascular disease; **TIA** = transient ischaemic attack.

For all candidate cases, enrolment records were assessed to identify and exclude patients who were not continuously enrolled in the health plan for the 365-day 'pre-event' period leading up to the index date. The claims histories for each candidate case were scanned during the pre-event period to identify CVD history (MI, stroke, TIA, CHF, angina pectoris) and selected co-morbidities of interest (hypertension, dyslipidaemias, type 2 diabetes). Indicator variables were also created to denote the presence or absence of non-cardiovascular conditions within broad disease categories defined by ICD-9-CM chapter heading (i.e. neoplasms, ICD-9-CM 140–239; blood disorders, 280–289; mental disorders, 290–319; circulatory system, 390–459; respiratory system, 460–519; and musculoskeletal system, 710–739).

A similar procedure was undertaken for patients who did not have evidence of a study event during calendar year 2003. For these candidate controls, an artificial 'index date' was randomly selected from this period and those not continuously

enrolled during the 365-day period leading up to this date were excluded. The final set of controls were selected using propensity score matching methods, which reduce selection bias by matching cases and controls based on their 'propensity' to fall into a predefined group.[21,22] In this case, the propensity to have each study event was estimated for each case and candidate control using logistic regression on a predefined set of demographic and clinical characteristics (age; sex; geographic region; indicators for co-morbid hypertension, dyslipidaemia and type 2 diabetes; indicators for history of MI, stroke, TIA, CHF, angina, PVD and atrial fibrillation; and indicators for a history of non-cardiovascular conditions grouped by ICD-9-CM chapter heading). Each case was then matched to one or more controls based on their propensity score using 'nearest neighbour' matching.[23] For events with sample sizes <1000 patients, a 1 : 3 match of cases to controls was used; for events with sample sizes 1001–4000, a 1 : 2 match was employed; and for events with sample sizes >4000 patients, a 1 : 1 match was utilized.

We included each case and control so long as they had not disenroled from the health plan for any reason other than death. Deaths were retained in the time series to reflect the impact of event-related mortality on healthcare costs over time. The only exception to this was for cardiac arrest and fatal MI and stroke events, which, by definition, only have 30 days of follow-up data.

The same claims database was utilized to construct the sample for the post-3 year cost estimation; however, in this analysis, attention was focused on calendar year 2002. Healthcare claims were scanned during 2002 to identify patients who had any of the non-fatal events of interest. As in the acute-phase analysis, the earliest observed event for each patient was deemed his or her one study event, and in cases in which multiple events occurred simultaneously, the event was included as a case for the procedure rather than the diagnosis. The date of the event was defined as the patient's index date, and for candidate cases, enrolment records were scanned to identify those who survived and were continuously enrolled in the plan for at least 3 years following the index date. Those not meeting these criteria were excluded. Additional scanning of each remaining patient's healthcare claims in 2002 was performed to identify the co-morbidities of interest and disease history. A similar exercise was undertaken for patients who were continuously enrolled in the health plan and did not experience any of the study events during calendar year 2002; for these candidate controls, an artificial index date was randomly selected from calendar year 2002. As in the acute-phase cost estimation, the final set of controls was selected based on propensity score matching.

Due to small sample sizes in the post-year 3 analysis, events were collapsed as follows: CHD (includes MI, angina, cardiac arrest), stroke (includes haemorrhagic and ischaemic strokes) and PCI (percutaneous coronary intervention) [includes CABG and PTCA with and without stent]. PVD, TIA and CHF all had adequate sample sizes for this analysis.

Econometric Analysis

Econometric analyses of total cost (including inpatient, outpatient and prescription drug services) were performed using separate cost prediction equations for cases and controls. The cost prediction equations took the form of equation 1:

$$Y = \beta_0 + \beta_i X_i + \beta_j T_j + \varepsilon \qquad \text{(Eq. 1)}$$

where Y is the total monthly cost, X_i is a vector of cross-sectional explanatory variables (age, sex, disease history, co-morbid conditions), T_j is a vector of dummy variables representing month-past index date, β is the beta coefficient to be estimated and ε is the error term. The monthly cost variable reflects amounts paid by the health plan and patient co-pays for all healthcare claims for institutional, provider and pharmacy services provided during that month (including both cardiovascular and non-cardiovascular care). For claims that represent services spanning more than one time period (e.g. hospitalization on day 28 of one month and discharge on day 8 of the next), we used an apportioning algorithm to divide the costs among the relevant months.

Table II. Sample sizes and selected characteristics for acute-phase cases and their matched controls

Event	Cases	Controls
Non-fatal MI	5106	5106
mean age (y)	65.0	65.4
% female	38.4	32.6
Resuscitated cardiac arrest	200	600
mean age (y)	62.2	61.7
% female	49.0	47.3
Hospitalization for angina	8946	8946
mean age (y)	60.0	60.5
% female	43.9	41.4
Hospitalization for CHF	12 629	12 629
mean age (y)	70.0	71.3
% female	51.2	49.5
Non-fatal ischaemic stroke	6821	6821
mean age (y)	68.4	69.9
% female	51.8	49.5
Non-fatal haemorrhagic stroke	1003	2006
mean age (y)	59.8	62.2
% female	50.5	50.0
Transient ischaemic attack	18 245	18 245
mean age (y)	64.3	65.5
% female	54.9	55.2
Hospitalization for PVD	6810	6810
mean age (y)	66.5	68.4
% female	43.5	44.9
CABG	983	2949
mean age (y)	62.4	63.4
% female	18.5	17.8
PTCA without stent	1879	3758
mean age (y)	60.1	61.5
% female	27.1	25.0
PTCA with stent	9398	9398
mean age (y)	60.4	61.8
% female	25.2	22.7
Fatal myocardial infarction	466	1398
mean age (y)	73.0	74.3
% female	42.7	44.3
Fatal cardiac arrest	246	738
mean age (y)	64.0	66.1
% female	44.3	48.1
Fatal ischaemic stroke	434	1302
mean age (y)	73.0	74.2
% female	49.8	45.5
Fatal haemorrhagic stroke	233	699
mean age (y)	64.6	66.5
% female	48.1	47.5

CABG = coronary artery bypass graft; **CHF** = congestive heart failure; **PTCA** = percutaneous transluminal coronary angioplasty; **PVD** = peripheral vascular disease.

The cost prediction equations were estimated using generalized linear models (GLMs), which allow for the distribution of the dependent variable to be non-Normal; in addition, they allow for the effect of the predictors on the dependent variable to be something other than linear.[24,25] Based on results of specification tests, models with a Poisson distribution and log link were selected for the GLM estimation (i.e. Poisson error structure).

The analytic datasets in the post-year 3 analysis were cross-sectional in nature, but did not require time series. Instead, an average annual cost was calculated for cases and controls using the total costs of care during the follow-up period. Analyses of each patient's healthcare claims were undertaken for the period beginning 3 years past the index date and continuing to the patient's 'end date', defined as the date of the patient's disenrolment from the plan or end of the study database, whichever occurred first. For each case and control, an annual cost of care variable was defined as total cost incurred between index date and end date divided by the total number of years of follow-up (equation 2):

$$= \frac{total\ cost}{(end\ date - index\ date)/365} \qquad \text{(Eq. 2)}$$

This estimate was combined with post-event survival probabilities to project mean annual costs for the patient's remaining years of life.

An equation similar to that estimated in the acute-phase analysis was utilized in the post-year 3 estimation (except that the latter did not include the longitudinal component). In both analyses, the differential cost between cases and controls was assumed to represent the attributable costs for the event. The cost prediction equations took

the form of equation 3:

$$Y = \beta_0 + \beta_i X_i + \varepsilon \quad \text{(Eq. 3)}$$

where Y is the annual cost variable defined above, X_i is a vector of cross-sectional explanatory variables (age, sex, disease history, co-morbid conditions), β is the beta coefficient to be estimated and ε is the error term.

Finally, we estimated the case and control equations for seven events (i.e. non-fatal MI, resuscitated cardiac arrest, angina, non-fatal ischaemic stroke, non-fatal haemorrhagic stroke, CABG and PTCA with stent) using the mean values of the patient characteristics for each sample. We estimated costs on a monthly basis in the acute period (i.e. the first 3 years after the event) and over a lifetime horizon. We inflated all cost data from the claims database to 2007 values using the medical care component of the Consumer Price Index. In the analyses, all future costs were discounted to present value using a 3% annual rate.

Results

Sample sizes and selected patient characteristics for each event are contained in tables II and III. In the acute-phase analysis, the most prevalent events were TIA, CHF, PTCA with stent and angina; in the post-year 3 analysis, TIA and the composite categories of CHD and PCI were the most prevalent. For the acute-phase analysis events, mean age ranged from 60 years among those hospitalized for angina, non-fatal haemorrhagic stroke and PTCA (with or without stent) to 73 years among patients with fatal MI and fatal ischaemic stroke. In general, patients were only slightly less likely to be female than male, with the exception of non-fatal MI (38% female), PVD (44% female), CABG (19% female), PTCA without stent (27% female) and PTCA with stent (25% female). Complete summaries of sample characteristics, along with beta coefficients, standard errors and p-values from the regression equations are contained in the Supplemental Digital Content.

Patients with non-fatal MI incurred significant costs in the month in which the event occurred (figure 1a), and maintained higher cumulative costs relative to their matched controls throughout the remainder of their lifetime. In the month of the event, attributable costs (i.e. total costs for cases minus total costs for controls) were about $US34 200. The cumulative cost differential between cases and controls continued to climb in the years immediately following the event, with attributable 12-month costs totalling $US55 500, attributable 24-month costs totalling $US64 900 and attributable 36-month costs totalling $US73 300. Lifetime attributable costs for non-fatal MI were $US55 400.

Patients experiencing a cardiac arrest from which they were resuscitated had higher costs in the month of the event relative to their matched controls; furthermore, the cost difference continued to increase over the 36-month period following the event (figure 1b). While the attributable cost of resuscitated cardiac arrest was $US34 500 in the month of the event, at the end of 12 months

Table III. Sample sizes and selected characteristics for post-year 3 cases and their matched controls

Event	Cases	Controls
Coronary heart disease[a]	4669	4669
mean age (y)	62.7	63.3
% female	43.7	39.6
Hospitalization for CHF	1820	3640
mean age (y)	67.2	68.7
% female	54.1	49.5
Stroke[b]	1799	3598
mean age (y)	67.1	67.9
% female	52.6	50.9
Transient ischaemic attack	6125	6125
mean age (y)	65.3	66.3
% female	56.6	56.0
Hospitalization for PVD	1590	3180
mean age (y)	66.4	68.0
% female	45.6	45.3
Percutaneous coronary intervention[c]	3231	6462
mean age (y)	61.8	62.6
% female	27.9	23.6

a Includes resuscitated cardiac arrest, hospitalization for angina and non-fatal myocardial infarction.

b Includes non-fatal ischaemic and non-fatal haemorrhagic stroke.

c Includes PTCA with (any) stent, PTCA without stent and CABG.

CABG = coronary artery bypass graft; **CHF** = congestive heart failure; **PTCA** = percutaneous transluminal coronary angioplasty; **PVD** = peripheral vascular disease.

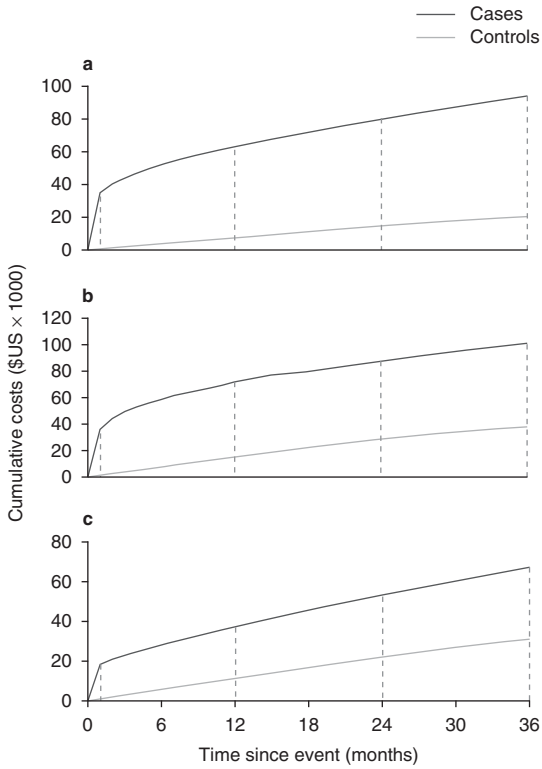

Fig. 1. Cumulative 3-year costs (year 2007 values) for (**a**) myocardial infarction cases and controls; (**b**) resuscitated cardiac arrest cases and controls; and (**c**) angina cases and controls.

the cost differential was $US56 500, at the end of 24 months it was $US58 700 and at the end of 36 months it was $US62 900. The lifetime cumulative cost difference between cases and controls was about $US72 200.

Attributable costs for hospitalization for angina (without coronary revascularization) were about $US17 300 in the month of the index hospitalization, $US26 000 at the end of 12 months, $US31 100 at the end of 24 months and $US36 000 at the end of 36 months (figure 1c). Attributable costs for angina were about $US32 700 in the last year of life.

As illustrated in figure 2a, patients who experienced an ischaemic stroke incurred higher costs in the month in which the stroke occurred relative to controls, with an attributable cost of about $US10 600. The cost differential continued to rise over 3 years post-event, after which it decreased. While the attributable cost of ischaemic stroke was $US18 300 at 12 months, $US18 800 at 24 months and $US20 000 at 36 months, it was $US4000 over a lifetime horizon.

Attributable costs of non-fatal haemorrhagic stroke were about $US39 400 in the month in which the event occurred and $US65 500 at 1 year post-event (figure 2b). The cost differential between cases of haemorrhagic stroke and controls increased further to $US69 000 at the end of 24 months and $US71 600 after 36 months, before levelling off somewhat. The total lifetime attributable cost of haemorrhagic stroke was $US50 600.

As can be seen in figure 2c, the difference in costs between CABG cases and controls was highest at the end of the first year following the event. In the month of the event, attributable costs were about $US31 900; at the end of the first year, this figure was $US32 900. The cost differentials at 24 and 36 months were $US29 900 and $US28 700, respectively. Attributable lifetime costs for CABG were about $US21 300.

The difference in costs between PTCA stent cases and controls was $US22 200 in the month of the surgery; it increased slowly thereafter (figure 2d). At the end of 12 months, attributable costs were $US31 000; corresponding cost figures for 24 and 36 months were $US35 900 and $US36 000, respectively. The attributable cost over a lifetime horizon was $US40 200.

Finally, attributable costs for fatal events were $US15 400 for MI, $US15 300 for cardiac arrest, $US9500 for ischaemic stroke and $US15 400 for haemorrhagic stroke (data not shown).

Discussion

We developed cost prediction equations to estimate the costs of several different cardiovascular events over various time horizons for patients with specific demographic and clinical characteristics. In this paper, we have described the methods and data used to construct the equations, and provided results for seven events among patients of given characteristics (i.e. using the mean values of our study samples).

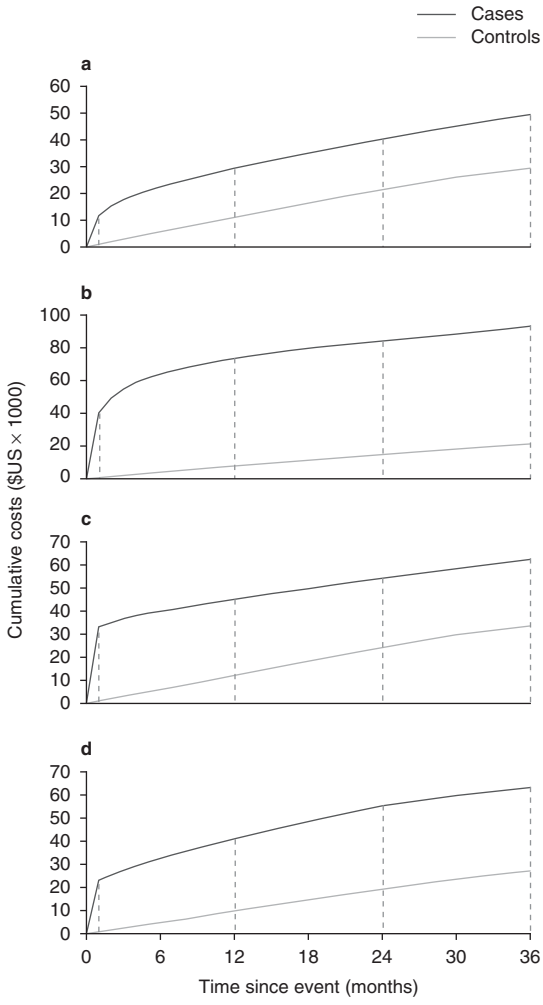

Fig. 2. Cumulative 3-year costs (year 2007 values) for (**a**) ischaemic stroke cases and controls; (**b**) haemorrhagic stroke cases and controls; (**c**) coronary artery bypass graft cases and controls; and (**d**) percutaneous transluminal coronary angioplasty with stent cases and controls.

Previously published economic studies in CVD have ranged from incidence-based cost-of-illness studies[9-12] to economic evaluations of primary and secondary prevention interventions.[13-16] Both types of analyses require incidence-based cost estimates of the clinical endpoints of interest (e.g. MI, stroke) for estimation of economic impact of the intervention under study. Costs have been obtained from published literature or an administrative healthcare claims database. Existing studies are somewhat outdated; moreover, no other study has estimated costs for the wide range of events included here (15 cardiovascular events in total). Finally, the costs of CVD events can vary substantially by patient demographic and clinical characteristics. Our analysis provides up-to-date cost estimates based on an expansive list of patient and clinical characteristics over various lengths of follow-up ranging from 1 month to lifetime. Results from our study will be especially useful in estimating costs of cardiovascular endpoints included in piggyback economic evaluations as well as those included in cost-effectiveness analyses of primary and secondary prevention interventions. Our methods allow for results to be updated as treatment patterns change and medical technology evolves.

Our study is not without limitations, largely related to the retrospective use of administrative claims data. First, the study sample was drawn from a database of enrolees in a large US health plan, and therefore may not be representative of the US population as a whole. In particular, the Medicare Advantage enrolees (those aged >65 years) in the plan may be healthier than their counterparts not enrolled in Medicare Advantage plans. However, the database used in this study had the benefits of yielding large sample sizes for most of the CVD events, allowing us to estimate costs beyond 3 years, and containing a plan population that is generally geographically diverse across the US, while at the same time containing a sample of Medicare beneficiaries, a population in which CVD is most prevalent. Note that our estimates are not expected to be generalizable to other countries, due to differences in payment systems, treatment patterns, costs and the like.

Other limitations associated with the use of administrative data relate to information that is not available from claims. Ideally, the cost prediction equations would include all patient characteristics and behaviours that may affect costs for patients with CVD. Two obvious omissions are smoking habit and body mass index. Smoking and obesity have been shown to be associated with the development of CVD;[1] they are also associated with increased healthcare costs.[26]

However, because the claims database does not contain information on smoking behaviour or height and weight, we were unable to include these measures in the cost prediction equations. In addition, claims databases do not capture information regarding the severity of co-morbid conditions, which may be important determinants of outcomes and costs. It is unknown what biases are present in our results due to these and other omissions.

Claims databases also have limited information on death. Specifically, the only deaths that can be tracked in the study database are those that occurred in hospital. While patients who died were retained in the acute-phase analysis to reflect the impact of event-related mortality on healthcare costs over time, we acknowledge that deaths are likely to be unrepresented in the analysis. If our analysis excludes most deaths, our cost estimates are likely to be biased upward, as they are reflective of survivor costs only. However, if health plan members who died were not formally disenroled from the health plan, their records would have been included in our analysis each month as zero costs, and estimates would be closer to true event-related costs. The effect of unknown mortality therefore is not quantifiable.

The study database contains healthcare claims for physician, hospital and prescription drugs, but does not provide information on costs for long-term care services such as those received in a rehabilitation facility or nursing home. This means that our cost results for events that require these types of services – most notably stroke – are underestimated. Should researchers use results of our study to estimate the lifetime costs of stroke, they should be cognizant of the fact that costs for long-term rehabilitation and nursing home care are not included. Indirect costs such as those for lost productivity are also not contained in claims databases. The study results are therefore reflective of direct medical care costs only.

It also should be noted that we did not track events that may have occurred after the index event. Patients who have a cardiovascular event are at risk of experiencing subsequent events over their lifetime. Our equations cannot be used to estimate the costs of two events occurring in one patient, such as an MI and a revascularization procedure, or a subsequent event occurring years after the initial event, such as an MI followed by CHF. Although it was beyond the scope of this study to track multiple or subsequent events, the cumulative lifetime cost estimates presented in this study reflect the risks and costs of subsequent events occurring in these patients.

Our cost equations allow for the estimation of costs over various time horizons, including a lifetime. In general, it is challenging to follow patients for more than 3 years within the context of a claims database because of disenrolment patterns. For that reason, we used different methods to estimate acute (i.e. 3-year) costs versus those beyond 3 years. While we estimated monthly costs over 36 months in the acute period, for years 4 through the end of the follow-up period (i.e. the patient's disenrolment in the plan), we estimated an average annual cost and assumed that this annual cost would apply for the remaining years of the patient's lifetime. While average cost estimates after year 3 may not be precise, we believe that, on average, attributable costs of events are captured in the cost differences between cases and controls.

The average annual costs estimated in the post-year 3 analysis were multiplied by event-specific survival probabilities to account for the fact that our post-year 3 sample was limited to patients who were alive during that period. Based on clinical judgement, we assumed that the excess risk of death associated with cardiovascular events persists for 10 years following the event, after which patients revert to age- and sex-specific risks of death consistent with that of the general population. Event-specific excess hazards for death were applied to background hazards for the US population to arrive at event-specific survival probabilities. Data on hazards for death among those with versus those without events of interest were obtained from three community-based cohort studies; to the extent that patients in these studies differ from our patient population, event-specific mortality may be biased. However, to the best of our knowledge, these studies represent the most relevant data on long-term risks of death associated with CV events to date. In addition to generalizability, as a substantial proportion of

the US background HRs of death are attributable to CVD, we acknowledge that our approach may overestimate the risk of death from events during years 4–10.

Conclusion

This study provides updated estimates of the acute and long-term costs associated with various cardiovascular events in the US. Results from this study should be a valuable resource to researchers conducting pharmacoeconomic and outcomes research studies in CVD in the years to come.

Acknowledgements

This study was sponsored by Pfizer Inc. Amy O'Sullivan, Jaime Rubin, Joshua Nyambose and David Thompson were employees of i3 Innovus, Medford, MA, USA at the time this research was conducted; i3 Innovus was a paid consultant to Pfizer Inc in connection with the development of the manuscript. David Cohen and Eric Peterson were paid consultants to Pfizer Inc in connection with the development of this manuscript. Andreas Kuznik is a full-time employee of Pfizer Inc. The authors would like to thank Eric Peterson for his contributions to this study. David Cohen has received consulting fees or speaking honoraria from Eli Lilly, Daiichi Sankyo, Medtronic, Cordis, Volcano and St. Jude Medical. He currently receives research grant support from Boston Scientific, Abbott Vascular, Medtronic, Edwards Lifesciences, Astra Zeneca and Medrad.

References

1. Framingham Heart Study. A project of the National Heart, Lung, and Blood Institute and Boston University [online]. Available from URL: http://www.framinghamheartstudy.org [Accessed 2008 Nov 1]
2. Anderson KM, Odell PM, Wilson PWF, et al. Cardiovascular disease profiles. Am Heart J 1990; 121: 293-8
3. Anderson KM, Wilson PWF, Odell PM, et al. An updated coronary risk profile: a statement for health professionals. Circulation 1991; 83 (1): 356-62
4. D'Agostino RB, Russell MW, Huse DM, et al. Primary and subsequent coronary risk appraisal: new results from the Framingham Heart Study. Am Heart J 2000 Feb; 139: 272-81
5. Weinstein MC, Coxson PG, Williams LW, et al. Forecasting coronary heart disease incidence, mortality, and cost: the Coronary Heart Disease Policy Model. Am J Public Health 1987; 77: 1417-26
6. Grover SA, Abrahamowicz M, Joseph L, et al. The benefits of treating hypercholesterolemia to prevent coronary heart disease: estimating changes in life expectancy and morbidity. JAMA 1992; 267: 816-22
7. Babad H, Sanderson C, Naidoo B, et al. The development of a simulation model of primary prevention strategies for coronary heart disease. Health Care Manag Sci 2002; 5: 269-74
8. Cooper K, Davies R, Roderick P, et al. The development of a simulation model of the treatment of coronary heart disease. Health Care Manag Sci 2002; 5: 259-67
9. Birnbaum H, Leong S, Kabra A. Lifetime medical costs for women: cardiovascular disease, diabetes, and stress urinary incontinence. Womens Health Issues 2003; 13 (6): 204-13
10. Russell MW, Huse DM, Drowns S, et al. Direct medical costs of coronary artery disease in the US. Am J Cardiol 1998; 81: 1110-5
11. Taylor TN, Davis PH, Torner JC, et al. Lifetime cost of stroke in the US. Stroke 1996; 27: 1459-66
12. Lee WC, Christensen MC, Joshi AV, et al. Long-term costs of stroke subtypes among Medicare beneficiaries. Cerebrovasc Dis 2007; 23: 57-65
13. Mark DB, Knight JD, Cowper PA, et al. Long-term economic outcomes associated with intensive versus moderate lipid-lowering therapy in coronary artery disease: results from the Treating New Targets (TNT) trial. Am Heart J 2008; 156: 698-705
14. Miller PS, Smith DG, Jones P. Cost effectiveness of rosuvastatin in treating patients to low-density lipoprotein cholesterol goals compared with atorvastatin, pravastatin, and simvastatin (a US Analysis of the STELLAR Trial). Am J Cardiol 2005; 95: 1314-9
15. Taylor DC, Pandya A, Thompson D, et al. Cost-effectiveness of intensive atorvastatin therapy in secondary cardiovascular prevention in the United Kingdom, Spain, and Germany, based on the Treating New Targets Study. Eur J Health Econ 2009 Jul; 10: 255-65
16. Ramsey SD, Clarke LD, Roberts CS, et al. An economic evaluation of atorvastatin for primary prevention of cardiovascular events in type 2 diabetes. Pharmacoeconomics 2008; 26: 329-39
17. Hartunian NS, Smart CN, Thompson MS. The incidence and economic costs of cancer, motor vehicle injuries, coronary heart disease, and stroke: a comparative analysis. Am J Public Health 1980; 70: 1249-60
18. Dennis MS, Burn JPS, Sandercock PAG, et al. Long-term survival after first-ever stroke: the Oxfordshire Community Stroke Project. Stroke 1993; 24: 796-800
19. Lampe FC, Whincup PH, Wannamethee SG, et al. The natural history of prevalent ischaemic heart disease in middle-aged men. Eur Heart J 2000; 21: 1052-62
20. Mosterd A, Cost B, Hoes AW, et al. The prognosis of heart failure in the general population. Eur Heart J 2001; 22: 1318-27
21. Rosenbaum P, Rubin D. The central role of the propensity score in observational studies for causal effects. Biometrika 1993; 70: 41-55
22. D'Agostino Jr RB. Tutorial in biostatistics: propensity score methods for bias reduction in comparison of a treatment to a non-randomized control group. Stat Med 1998; 17: 2265-81
23. Coca-Perraillon M. Local and global optimal propensity score matching [paper 185-2007]. SAS Global Forum, 2007 [online]. Available from URL: http://www.biostat.umn.edu/~will/6470stuff/Class25-10/SAS%20matching%202007.pdf [Accessed 2011 Mar 24]

24. Mullahy J. Much ado about two: reconsidering retransformation and two-part models in health econometrics. J Health Econ 1998; 17: 247-81

25. Buntin MB, Zaslavsky AM. Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. J Health Econ 2004; 23: 525-42

26. Kruse M, Davidsen M, Madsen M, et al. Costs of heart disease and risk behaviour: implications for expenditure on prevention. Scand J Public Health 2008; 36: 850-6

Correspondence: Dr *Amy K. O'Sullivan*, i3 Innovus, 10 Cabot Road, Suite 304, Medford, MA 02155, USA.
E-mail: amy.osullivan@i3innovus.com