Appendix 1 (as supplied by the authors): Methods and supplemental tables

Effects of Cardiovascular and Cerebrovascular Health Events on Work and Earnings: A Population-Based, Retrospective Cohort Study

Allan Garland, MD, MA , Sung-Hee Jeon, PhD, Michael Stepner, BA, Michelle Rotermann, MA, Randy Fransoo, PhD, Hannah Wunsch, MD, MSc, Damon C. Scales, MD, PhD, Theodore J. Iwashyna, MD, PhD, Claudia Sanmartin, PhD

CONTENTS

METHODS: DATASETS AND LINKAGE.

The data for this retrospective, population-based cohort study came from Canadian Hospitalization and Taxation Database (C-HAT), a previously described linkage of population-based Canadian hospital and tax data.[1] Hospital data in C-HAT derive from the Discharge Abstract Database (DAD),[2] which contains detailed information for all acute care hospitalizations in Canada, excepting Quebec. It includes: (a) up to 25 diagnoses coded in International Statistical Classification of Diseases (ICD)-10-CA format, each accompanied by a code identifying whether it occurred before or after hospital admission, and, (b) up to 20 in-hospital procedures coded by the Canadian Classification of Health Interventions. Tax data in C-HAT derive from the T1 Family File (T1FF), containing yearly tax returns for all Canadians. The hospital and tax data are provided to Statistics Canada on a regular basis by the Canadian Institute for Health Information and the Canadian Revenue Agency, respectively.

As described elsewhere, the hospital and tax data were linked using a deterministic match on sex, date of birth, and postal code.[1] We successfully linked 88.4% of DAD hospitalization records (n=48,203,982) to tax data, with internal validation indicating an overall true-positive linkage rate of 97.8%. For this study we used C-HAT excluding the three territories, where the linkage rate was lower than in the provinces.[1]

METHODS: STUDY COHORTS.

Separate analyses were performed for AMI, cardiac arrest and stroke, using matched cohorts of exposed and unexposed control individuals, where exposures were hospitalization for these health events. We use the term *index year* to refer to the calendar year of the health event event for exposed, and the matched year for unexposed; for simplicity we refer to years before and after the index year for both exposed and unexposed as pre-event and post-event, respectively. We indicate the index year as Y0, the third year following it as Y+3, etc.

Based on sample size analysis (Appendix p. 4), we used the DAD to identify individuals admitted to an acute care hospital for AMI, cardiac arrest or stroke during calendar years 2008-2010. Because of lower yearly incidence in cardiac arrest, for that event we added three additional years, 2005-2007, excluding Manitoba for 2005-2007 due to incomplete data.

Inclusion criteria for exposed were: age 40-61 years in Y0; identified in tax data in Y-2, Y-1 and Y+3; working in Y-2 and Y-1 as indicated by non-zero earnings; hospitalization with AMI, cardiac arrest or stroke in Y0; and alive at the end of Y+3. ICD-10 diagnostic codes (Appendix eTable 1) were chosen to maximize positive predictive value, by excluding codes that may not represent the entities of interest. We excluded individuals who had experienced the health event of interest during the three years preceding the index event to ensure that labor market outcomes were attributable to the index health event rather than a prior one (Appendix eTable 1). We restricted the age range because AMI, cardiac arrest and stroke are rare below age 40 years, and to provide at least 3 full years of follow-up before the most common retirement age of 65.

Correctly assessing the index hospital length of stay required taking account of inter-hospital transfers, which create separate DAD records. Accordingly, we combined individual records across inter-hospital transfers into complete episodes of hospital care;[3] these episodes were the units of measure for index hospitalizations.

Unexposed control subjects for each study year were chosen separately for the three conditions. They comprised individuals who were not hospitalized during Y0 for AMI, cardiac arrest and stroke. Because this pool was extremely large, they were chosen from a 30% random sample of the linked hospital-tax database. Unexposed subjects were required to satisfy the same criteria as the exposed cohorts, except they did not have the primary health event of interest in Y0. To more reliably establish the absence of the three health event conditions, we used an expanded set of diagnostic codes to exclude these conditions during the three years prior to Y0 (Appendix eTable 1).

We excluded extreme earners, defined as being in the top and bottom 0.25% of the earnings distribution of the population aged 37-64 from 2003 to 2013. All those excluded at the low end of the distribution had negative total earnings, indicating net losses from self-employment.

METHODS: MATCHING.

For this study, we matched candidate unexposed individuals to exposed individuals using Coarsened Exact Matching (CEM), a method that is distinct from usual exact matching.[4,5] CEM divides each matching variable into strata. For intrinsically categorical variables this stratification is obvious. For each metric variable, the entire observed range is subdivided into bins, or "coarsened". For K matching variables, after coarsening there exists a K-dimensional space of coarsened bins. All included subjects are assigned to their appropriate K-dimensional bin. Subjects assigned to the same coursened bin are considered to be identical for the purpose of the matching. Weights are assigned according to the number of cases and controls in each bin, to render the cohort of cases to be similar to the cohort of controls; bins with at least one one case and one control subject are retained. As described by Iacus,[4,5] CEM has numerous advantages over balancing approaches such as Propensity Scores: (a) it controls covariate imbalance, the degree of model dependence, the estimation error and bias in the causal quantity of interest; (b) most matching methods attempt to match only in expectation, approximating randomization, while CEM also balances locally and in sample, thus approximating the more efficient randomized block design; (c) other matching methods must be preceded by an entirely different algorithm limiting covariates to areas of common empirical support -- but CEM does this automatically as a natural part of the same matching algorithm; (d) it is approximately invariant to measurement error; (e) global multivariate differences are controllable by easy-to-understand local decisions about specific variables and their measurement characteristics; (f) it chooses the maximum imbalance in advance, for each matching variable separately; (g) the user is certain of the maximal level of multivariable imbalance that results; (h) it automatically eliminates observations beyond the parameter range of each matching variable that is common for both groups; (i) unlike many matching processes, including Propensity Scores, it does not violate the Congruence Principal; and (j) it bounds the estimation error in the final quantity of interest. In addition, for CEM, changes in maximum imbalance on one matching variable does not affect maximum imbalance on others, so that the recommended process of balance checking, tweaking, and repeatedly rerunning the matching is eliminated, as is the uncertainty about whether the matching procedure will improve balance at all.

METHODS: ANALYSIS.

We quantified the effect of health events on mean yearly earnings using matched difference-in-difference, ordinary least squares regression. This method compares the pre-event to post-event change in earnings of exposed with a health event to the analogous change of unexposed without the health event. This "double difference" represents the effect on earnings attributable to the health event;[6] since we matched exposed and unexposed on Y-1 earnings, this quantity is identical to the post-event difference between average earnings of the weighted exposed and unexposed cohorts. Although it is not well recognized in the medical literature, causal inference achieved by combining matching with parametric regression analysis is well-established in other areas.[6-8] Because of our large sample sizes, the Central Limit Theorem allowed us to compare the mean values between groups, regardless of whether earnings follow a Gaussian distribution in the population.[9] To compare proportions working post-event between exposed and unexposed, we used weighted multivariable probit regression on the matched cohorts.

We performed three prespecified subset analyses. First, for outcomes in Y+3, we restricted analyses to individuals who were working in Y+3, as indicated by nonzero earnings in that year. Second, to reduce confounding due to false-positive identification of health events, we restricted exposed to those whose index hospital lengths of stay were at least three days. Finally, to reduce the contribution of individuals whose health events led to early retirement, we restricted analyses to those individuals who were 40-55 years of age in Y0. We conducted heterogeneity analyses to assess whether the labor market effects of AMI and stroke differed across strata of 12 variables (see Online supplement).

Analysis was performed using Stata 14 (StataCorp, College Station, TX). P-values < 0.05 were considered statistically significant.

Appendix to: Garland A, Jeon S-H, Stepner M, et al. Effects of cardiovascular and cerebrovascular health events on work and earnings: a population-based retrospective cohort study. *CMAJ* 2018. doi: 10.1503/cmaj.181238.

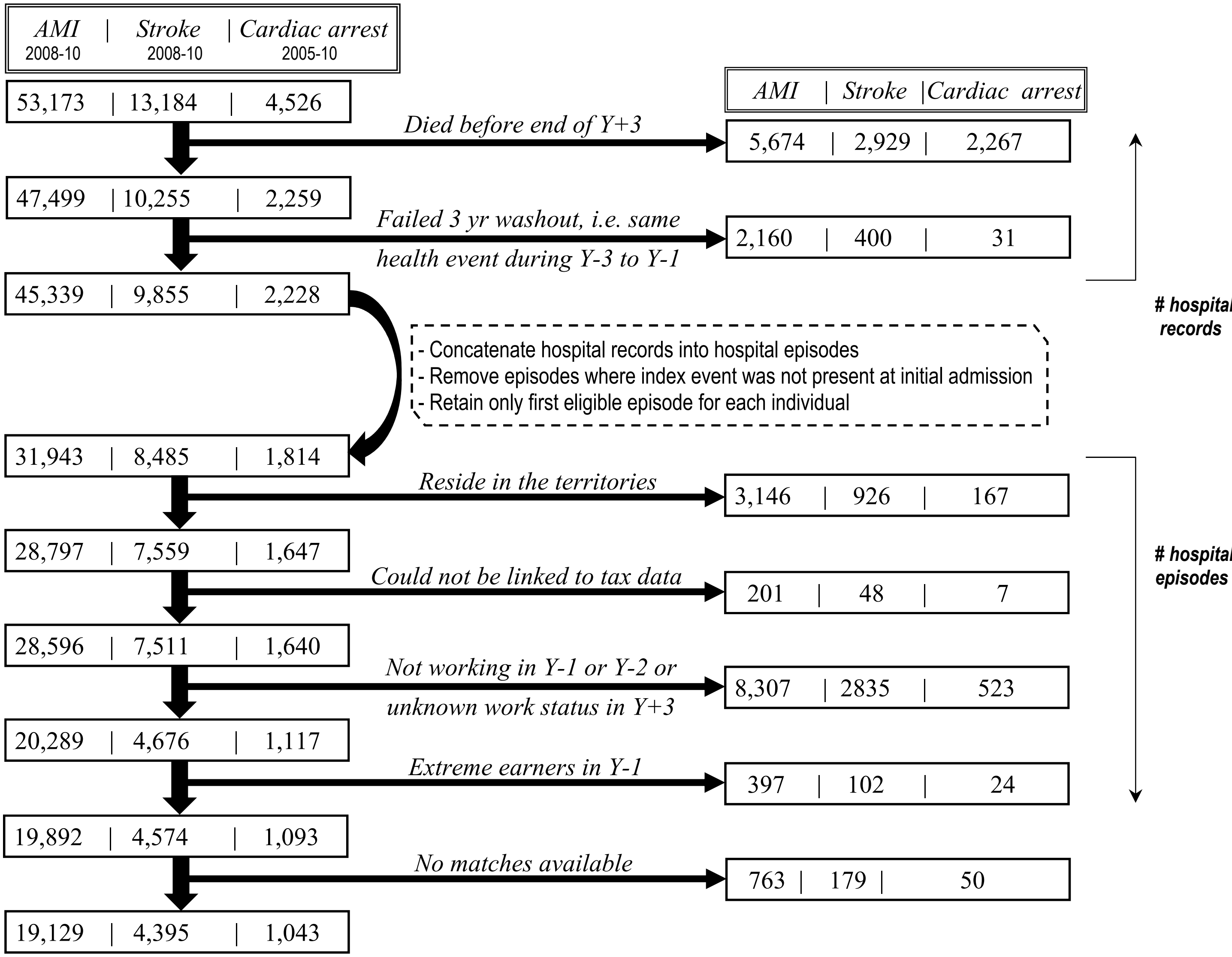
METHODS: HETEROGENEITY ANALYSIS.

We assessed whether the labor market effects of health events were heterogeneous by subcategories of 12 variables: age, sex, marital status, any self-employment income in Y-1, Y-1 earnings, burden of comorbid illness, index hospital length-of-stay (LOS), admission to an intensive care unit or stepdown unit, invasive mechanical ventilation, cardiac catheterization, any coronary artery revascularization procedure, and the subtype of stroke (ischemic, hemorrhagic). Comorbid illness was measured using the Charlson Comorbidity Index,[10] modified to exclude myocardial infarction for AMI and to exclude cerebrovascular disease for stroke. For each of these 12 analyses we: (a) stratified the variable of interest, (b) reweighted on a subset of the original matching variables in order to achieve balance across the strata, (c) performed matched difference-in-difference analysis for each stratum, and (d) compared across strata using Seemingly Unrelated Estimation.[11]. The key step was reweighting to create balance between the strata of the variable of interest,[12,13] while maintaining balance between the health event and matched, unexposed control cohorts effected by the coarsened exact matching weights. Because of progressively greater loss of cohort size with reweighting on more variables, we reweighted on 7 of the original 11 matching variables: year, age, sex, marital status in T-1, self-employment income in Y-1, total earnings in Y-1 and any non-obstetrical hospital days in Y-1 to Y-3; this led to acceptable loss of individuals (Appendix eTable 2). We did not perform heterogeneity analysis for cardiac arrest because even restricting to a minimal set of reweighting variables resulted in losses of up to 74% of those individuals.

METHODS: SAMPLE SIZE ESTIMATION.

Because this population-based analysis used all available data for each year, our sample size analysis amounted to calculating the number of years of data required for each of the three health events, acute myocardial infarction (AMI), cardiac arrest (cardiac arrest) and stroke. We assumed a Type I error rate of 1%, and a power of 90%. As we were unable to identify sample size analysis methods for difference-in-difference regression, we used a simplified, approximate approach, using PASS 2008 software (NCSS LLC, Kaysville, Utah) to calculate the power for improving the R^2 of multiple linear regression with 20 independent variables, by adding the variable of interest (i.e. whether the individual had an index health event). Even for an unfavorable scenario (increasing R^2 by only 0.01 from a low baseline of 0.20), this approach estimated the need for 1200 individuals for each health event. Based on data from the 2006/07 Canadian Discharge Abstract Database (DAD), we estimated usable yearly exposed cases, respectively, of 7000 for AMI, 200 for cardiac arrest and 1500 for stroke. This indicated that we use at least one year of data for AMI and stroke, and six years for cardiac arrest. Based on this, and to avoid singular years, we chose to use three years for AMI and stroke, and all six available years of DAD data for cardiac arrest. Since our tax data extended to the end of 2013, and our primary analysis required survival to three years after the health event, we used calendar years 2008-2010 for AMI and stroke, and 2005-2010 for cardiac arrest.

eFigure 1. Flow diagram of numbers of hospital records or episodes included in the analysis.



eTable 1. ICD-10 code inclusion and exclusion criteria for exposed health event and unexposed control cohorts.

Event	*Health event cohort inclusionary diagnoses [a]*	*Exclusionary diagnoses [b]*
Acute myocardial infarction	I21.x	I21.x, I22.x, I23.x
Stroke	I61.x, I63.x, I64.x	I61.x, I63.x, I64.x, I60.x, I62.x, G45.x, G46.x, H34.0, H34.1, I67.6
Cardiac arrest	I46.x, I49.0x	I46.x, I49.0x T81.8 = Other complications of procedures, not elsewhere classified I97.8 = Other post-procedural disorders of circulatory system, NEC

(a) Diagnosis type must indicate presence at hospital admission, i.e. Canadian Institute for Health Information Types 1, or 5, or M without also being listed as Type 2, or 6 without also being listed as Type 2.
(b) Any diagnosis type allowed. Presence of any of these excludes individuals from unexposed control cohorts, and provide washout criteria for both health event and unexposed control cohorts in the three years prior to the index year

eTable 2. Percentage of original health event cohorts lost in reweighting for heterogeneity analysis.

Heterogeneity variable assessed	*Health event cohort*		
	Acute myocardial infarction	*Stroke*	*Cardiac arrest*
Age	0.8	4.2	37.1
Sex	2.0	4.5	42.9
Marital status	1.3	6.6	47.5
±Self-employment in Y-1	4.7	21.6	55.1
Total earnings in the year prior to the health event (Y-1)	0.6	4.0	24.9
Charlson comorbidity	6.5	10.1	55.1
Index hospitalization length of stay	2.7	14.5	61.5
±Admitted to a special care unit during index hospitalization	1.9	12.6	74.3
±Mechanical ventilation during index hospitalization	7.1	38.5	42.8
±Cardiac catheterization during index hospitalization	2.3	--	51.8
±Coronary revascularization during index hospitalization	1.3	--	42.6
Stroke type	--	19.0	--

Reweighting was on strata of: year, age, sex, whether there was any self-employment income in Y-1, total earnings in Y-1, marital status in Y-1, and whether there were any non-obstetrical hospital days in Y-1 to Y-3.

eTable 3. Baseline characteristics for the exposed health event and unexposed control cohorts. (A) Myocardial Infarction, (B) Cardiac Arrest, (C) Stroke.

(A) Myocardial Infarction

Variable	*Before Matching*			*After Matching*		
	Health event cohort	*Potential controls*	*Standardized difference*	*Health event cohort*	*Potential controls*	*Standardized difference*
Number	19,892	5,055,533		19,129	1,820,644	
Age, Y0 (yrs; mean±SD)	53.0±5.5	49.8±6.0	0.558	53.0±5.5	52.9±5.5	0.012
Female (%)	19.2	49.9	-0.681	19.1	19.1	0.000
Province/region, Y-1 (%)						
Newfoundland	3.3	2.3	0.066	3.0	3.0	0.000
Prince of Edward Island	0.7	0.6	0.017	0.6	0.6	0.000
Nova Scotia	4.7	3.9	0.043	4.5	4.5	0.000
New Brunswick	4.8	3.2	0.080	4.6	4.6	0.000
Ontario	47.8	50.3	-0.049	49.3	49.3	0.000
Manitoba	5.6	4.7	0.041	5.3	5.3	0.000
Saskatchewan	4.2	4.0	0.008	3.9	3.9	0.000
Alberta	14.5	13.8	0.020	14.5	14.5	0.000
British Columbia	14.3	17.2	-0.081	14.3	14.3	0.000
Urban residency, Y-1 (%)	82.7	86.0	-0.091	83.9	83.9	0.000
Couple, Y-1 (%)	79.0	79.0	-0.001	79.9	79.9	0.000
Union member, Y-1 (%)	31.7	33.6	-0.040	31.1	31.1	0.000
Any self-employment, Y-1 (%)	16.8	17.2	-0.012	16.2	16.2	0.000
Earnings, Y-1 (2012$; mean±SD)	53,118± 44,435	54,600± 45,523	-0.033	53,263± 44,572	53,618± 45,883	-0.008
Earnings, Y-2 (2012$; mean±SD)	54,075± 44,329	54,423± 45,341	-0.008	54,115± 44,527	54,465± 45,954	-0.008
At least 1 hospital day, Y-3, Y-2, Y-1 (%)	12.1	7.7	0.146	10.6	10.6	0.000
Have children <18 yrs, Y-1 (%)	28.7	41.7	-0.275	29.0	29.9	-0.020
# of children <18 yrs, Y-1, of those with any mean±SD median (IQR)	1.6±0.8 1 (1-2)	1.7±0.8 1 (1-2)	-0.143	1.6±0.8 1 (1-2)	1.6±0.8 1 (1-2)	-0.020

Y0, calendar year of index health event; Y+3, third calendar year after Y0, etc.
Normalized difference = $(mean_2 - mean_1)/[(SD_1^2+SD_2^2)/2]^{1/2}$

Appendix to: Garland A, Jeon S-H, Stepner M, et al. Effects of cardiovascular and cerebrovascular health events on work and earnings: a population-based retrospective cohort study. *CMAJ* 2018. doi: 10.1503/cmaj.181238.

(B) Cardiac Arrest

Variable	*Before Matching*			*After Matching*		
	Health event cohort	*Potential controls*	*Standardized difference*	*Health event cohort*	*Potential controls*	*Standardized difference*
N	1,093	9,717,362		1,043	307,375	
Age, Y0 (yrs; mean±SD)	52.6±5.7	49.6±6.0	0.509	52.6±5.7	52.5±5.8	0.010
Female (%)	23.5	49.5	-0.561	23.7	23.7	0.000
Province, Y-1 (%)						
Newfoundland	3.5	2.3	0.069	3.2	3.2	0.000
Prince of Edward Island	0.4	0.6	-0.038	0.4	0.4	0.000
Nova Scotia	5.1	4.0	0.053	4.6	4.6	0.000
New Brunswick	3.8	3.4	0.026	3.7	3.7	0.000
Ontario	48.9	51.5	-0.050	50.6	50.6	0.000
Manitoba	2.8	2.5	0.023	2.5	2.5	0.000
Saskatchewan	5.0	4.1	0.043	4.7	4.7	0.000
Alberta	13.9	14.0	-0.004	13.8	13.8	0.000
British Columbia	16.5	17.5	-0.027	16.5	16.5	0.000
Urban residency, Y-1 (%)	89.3	86.0	0.101	90.1	90.1	0.000
Couple, Y-1 (%)	80.7	79.1	0.041	81.7	81.7	0.000
Union member, Y-1 (%)	33.6	33.9	-0.007	32.4	32.4	0.000
Any self-employment, Y-1 (%)	15.2	17.5	-0.063	14.8	14.8	0.000
Earnings, Y-1 (2012$; mean±SD)	54,941± 46,682	53,687± 44,746	0.027	55,192± 47,141	55,044± 47,162	0.003
Earnings, Y-2 (2012$; mean±SD)	54,923± 43,657	53,140± 44,173	0.041	54,827± 44,332	55,442± 46,679	-0.014
At least 1 hospital day, Y-3, Y-2, Y-1 (%)	21.0	8.2	0.367	19.1	19.1	0.000
Have children <18 yrs, Y-1 (%)	32.2	42.2	-0.207	32.5	31.8	0.015
# of children <18 yrs, Y-1, of those with any mean±SD median (IQR)	 1.6±0.8 1 (1-2)	 1.7±0.8 1 (1-2)	 -0.117	 1.6±0.8 1 (1-2)	 1.6±0.8 1 (1-2)	 0.002

Appendix to: Garland A, Jeon S-H, Stepner M, et al. Effects of cardiovascular and cerebrovascular health events on work and earnings: a population-based retrospective cohort study. *CMAJ* 2018. doi: 10.1503/cmaj.181238.

(C) Stroke

Variable	*Before Matching*			*After Matching*		
	Health event cohort	*Potential controls*	*Standardized difference*	*Health event cohort*	*Potential controls*	*Standardized difference*
N	4,574	5,072,227		4,395	888,481	
Age, Y0 (yrs; mean±SD)	53.1±5.8	49.8±6.0	0.554	53.1±5.8	53.0±5.8	0.012
Female (%)	34.7	49.7	-0.309	34.6	34.6	0.000
Province, Y-1 (%)						
Newfoundland	2.6	2.3	0.022	2.3	2.3	0.000
Prince of Edward Island	0.9	0.6	0.035	0.7	0.7	0.000
Nova Scotia	4.1	3.9	0.011	3.8	3.8	0.000
New Brunswick	4.2	3.2	0.051	3.9	3.9	0.000
Ontario	49.7	50.3	-0.011	51.3	51.3	0.000
Manitoba	4.9	4.7	0.011	4.6	4.6	0.000
Saskatchewan	4.2	4.0	0.009	3.9	3.9	0.000
Alberta	14.2	13.8	0.011	14.2	14.2	0.000
British Columbia	15.2	17.2	-0.055	15.4	15.4	0.000
Urban residency, Y-1 (%)	85.7	86.0	-0.008	86.9	86.9	0.000
Couple, Y-1 (%)	74.0	79.0	-0.118	74.8	74.8	0.000
Union member, Y-1 (%)	31.7	33.6	-0.041	31.1	31.1	0.000
Any self-employment, Y-1 (%)	15.2	17.2	-0.057	14.5	14.5	0.000
Earnings, Y-1 (2012$; mean±SD)	47,615± 41,746	54,585± 45,512	-0.160	47,657± 41,933	48,131± 43,884	-0.011
Earnings, Y-2 (2012$; mean±SD)	48,675± 41,906	54,422± 45,332	-0.132	48,585± 41,944	48,995± 43,693	-0.010
At least 1 hospital day, Y-3, Y-2, Y-1 (%)	17.3	8.0	0.283	15.7	15.7	0.000
Have children <18 yrs, Y-1 (%)	26.7	41.6	-0.320	26.9	28.6	-0.038
# of children <18 yrs, Y-1, of those with any mean±SD median (IQR)	1.6±0.8 1 (1-2)	1.7±0.8 2 (1-2)	-0.157	1.6±0.8 1 (1-2)	1.6±0.8 1 (1-2)	-0.055

Appendix to: Garland A, Jeon S-H, Stepner M, et al. Effects of cardiovascular and cerebrovascular health events on work and earnings: a population-based retrospective cohort study. *CMAJ* 2018. doi: 10.1503/cmaj.181238.

eTable 4. Temporal evolution of labor market outcomes for matched, weighted health event and unexposed control cohorts.

Year	Working (%)			Total annual earnings ($2012 CAD)			
	Health event cohort	Matched control cohort	Difference (health event minus control)	Health event cohort (mean±SD)	Matched control cohort (mean±SD)	Difference (health event minus control)	Percent difference [a]
ACUTE MYOCARDIAL INFARCTION							
Y-2	100.0	100.0		54,115±44,527	54,465±45,954		
Y-1	100.0	100.0		53,264±44,572	53,618±45,883		
Y0	95.3	96.3	-1.0	46,643±44,086	51,815±46,344	-4,817	-9.3
Y+1	89.4	93.1	-3.7	44,404±44,580	49,565±46,675	-4,807	-9.7
Y+2	86.4	91.1	-4.7	43,653±44,796	48,251±47,133	-4,244	-8.8
Y+3	85.1	90.2	-5.1	43,300±46,470	47,488±47,712	-3,834	-8.1
CARDIAC ARREST							
Y-2	100.0	100.0		54,827±44,332	55,443±46,679		
Y-1	100.0	100.0		55,192±47,141	55,044±47,162		
Y0	94.8	96.6	-1.8	46,036±46,840	53,390±47,524	-7,501	-14.1
Y+1	84.9	93.6	-8.7	41,833±48,237	51,718±48,169	-10,032	-19.4
Y+2	80.2	91.2	-11.0	40,468±46,956	50,022±48,380	-9,701	-19.4
Y+3	77.4	90.1	-12.7	37,891±45,709	48,887±49,421	-11,143	-22.8
STROKE							
Y-2	100.0	100.0		48,585±41,944	48,995±43,693		
Y-1	100.0	100.0		47,657±41,933	48,132±43,884		
Y0	91.9	95.5	-3.6	38,510±42,260	46,666±44,089	-7,681	-16.5
Y+1	75.5	92.0	-16.5	30,778±40,730	44,727±44,406	-13,474	-30.1
Y+2	-70.5	-89.7	-19.2	29,482±39,960	43,408±44,488	-13,451	-31.0
Y+3	69.0	88.8	-19.8	28,818±39,583	42,571±44,718	-13,278	-31.2

Y0, calendar year of index health event; Y+3, third calendar year after Y0, etc.
[a] Calculated as Difference column divided by Matched control cohort column

eTable 5. Changes in labor market outcomes in years after health events versus the year prior to health events attributable to the health events. Values are mean ± standard error.

Year	*Difference between health event and unexposed control cohorts in:*	
	Percent employed	*Pre-event to post-event change in annual earnings ($2012 CAD)*
	ACUTE MYOCARDIAL INFARCTION	
Y+1	-3.4 ± 0.2	-4,723 ± 210
Y+2	-4.4 ±0.2	-4,096 ± 229
Y+3	-5.0 ±0.3	-3,834 ± 249
	CARDIAC ARREST	
Y+1	-9.0 ±1.0	-10,159 ± 939
Y+2	-10.6 ±1.2	-9,534 ± 1,050
Y+3	-12.9 ± 1.3	-11,143 ± 1,113
	STROKE	
Y+1	-16.3 ±0.6	-13,406 ± 460
Y+2	-19.2 ±0.7	-13,515 ± 481
Y+3	-19.8 ± 0.7	-13,278 ± 498

Y+3 is the third calendar year after the year of the health event, etc.

Appendix to: Garland A, Jeon S-H, Stepner M, et al. Effects of cardiovascular and cerebrovascular health events on work and earnings: a population-based retrospective cohort study. *CMAJ* 2018. doi: 10.1503/cmaj.181238.

eTable 6. Absolute and relative results of stratified analyses of changes in labor market outcomes in the third year after health events (Y+3) versus the year prior to health events (Y-1) attributable to the health events. Each row represents a separate analysis.

Variable	*Total Yearly Earnings $2012 CAD mean ± standard error (as % of Y+3 earnings of unexposed controls)*	
	Acute myocardial infarction	*Stroke*
Age group		
40-49	-3,512±454 (6.4%)	-13,249±936 (27.1%)
50-55	-3,505±405 (7.0%)	-13,293±932 (29.6%)
56-61	-4,426±449 (11.2%)	-12,878±841 (37.1%)
Sex		
male	-3,727±267 (7.9%)	-13,500±609 (31.9%)
female	-3,520±967 (8.0%)	-12,257±1,068 (30.7%)
Y-1 earnings, terciles	***	***
low	-2,710±261 (13.4%)	-7,752±420 (43.4%)
middle	-4,099±324 (9.0%)	-13,871±680 (32.3%)
high	-6,070±952 (6.4%)	-23,780±2,133(26.4%)
Y-1 marital status		
single	-3,742±712 (8.2%)	-14,367±1,277 (35.8%)
married or common law	-3,957±271 (8.3%)	-12,720±568 (30.4%)
Any Y-1 self-employment		
no	-3,990±267 (8.4%)	-13,434±597 (33.0%)
yes	-3,300±923 (6.8%)	-11,753±1,777 (27.5%)
Charlson comorbidity index	***	***
0	-3,525±263 (7.5%)	-11,547±579 (27.5%)
≥1	-8,288±1,064 (17.5%)	-18,823±1,114 (43.3%)
Index hospital LOS, terciles	***	***
lowest	-2,623±383 (5.5%)	-7,362±891 (17.4%)
intermediate	-4,204±471 (8.8%)	-11,513±944 (27.5%)
highest	-5,291±515 (11.3%)	-22,374±1,101 (54.3%)
Special care unit		***
no	-3,703 ±533 (7.9%)	-11,636±606 (27.0%)
yes	-3,915±286 (8.2%)	-18,356±1,236 (43.3%)
Mechanical ventilation	***	***
no	-3,629±270 (7.6%)	-12,470±659 (29.2%)
yes	-7,572±962 (15.9%)	-26,198±2,659 (62.9%)
Cardiac catheterization		
no	-4,034±583 (8.5%)	not applicable
yes	-3,743±281 (7.9%)	
Revascularization		
no	-3,434±449 (7.2%)	not applicable
yes	-3,926±312 (8.3%)	
Stroke type		
ischemic	not applicable	-13,823±746 (31.6%)
hemorrhagic		-15,910±1,519 (36.9%)

*** $p<0.001$; ** $p<0.01$; * $p<0.05$ -- among the categories, by Seemingly Unrelated Estimation
LOS, length of stay

ONLINE SUPPLMENT REFERENCES

1. Sanmartin C, Reicker A, Dasylva A, et al. Data Resource Profile: Canadian Hospitalization and Taxation Database (C-HAT). *International Journal of Epidemiology.* 2018;47:687–g.
2. Discharge Abstract Database (DAD) Metadata. https://www.cihi.ca/en/discharge-abstract-database-metadata. Accessed August 2, 2018.
3. Fransoo R, Yogendran M, Olafson K, Ramsey C, McGowan K-L, Garland A. Constructing Episodes of Inpatient Care: Data Infrastructure for Population-based Research. *BMC Medical Research Methodology.* 2012;12:133.
4. Iacus SM, King G, Porro G. Causal inference without balance checking: Coarsened Exact Matching. Political Analysis. *Political Analysis.* 2012;20:1-24.
5. Iacus SM, King G, Porro G. Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association.* 2011;106:345-61.
6. Lechner M. The Estimation of Causal Effects by Difference-in-Difference Methods. *Foundations and Trends in Econometrics.* 2011;4:165-224.
7. Imbens GW, Rubin DB. *Causal inference for statistics, social, and biomedical sciences: an introduction*. New York: Cambridge University Press; 2015.
8. Ho D, Imai K, King G, Stuart EA. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis.* 2007;15:199-236.
9. Kleinbaum DG, Kupper LL, Muller KE, Nizam A. Applied Regression Analysis and Multivariable Methods. 3rd ed. Stamford: Brooks/Cole; 1998:15.
10. Quan H, Sundararajan V, Halfon P, et al. Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data. *Medical Care.* 2005;43:1130-9.
11. Wessie HM. Seemingly unrelated estimation and the cluster-adjusted sandwich estimator. *Stata Technical Bulletin.* 1999:34-47.
12. DiNardo J, Fortin NM, Lemieux T. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica.* 1996;64:1001-44.
13. Yagan D. Capital Tax Reform and the Real Economy: The Effects of the 2003 Dividend Tax Cut. *American Economic Review.* 2015;105:3531-63.

