

Open Access

## MyRisk\_Stroke Calculator: A Personalized Stroke Risk Assessment Tool for the General Population

Lisa Nobel,<sup>a,b</sup> Nancy E. Mayo,<sup>a,b</sup> James Hanley,<sup>a,b</sup> Lyne Nadeau,<sup>b</sup> Stella S. Daskalopoulou<sup>c</sup>

<sup>a</sup>Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada

<sup>b</sup>Division of Clinical Epidemiology, Department of Internal Medicine,

<sup>c</sup>Division of Internal Medicine, Department of Medicine, McGill University Health Centre, McGill University, Montreal, Quebec, Canada

Received October 19, 2012

Revised June 19, 2013

Accepted June 19, 2013

### Correspondence

Nancy E. Mayo, PhD  
Department of Medicine,  
McGill University Health Centre,  
Royal Victoria Hospital,  
Ross Pavilion R4.29,  
687 Pine Ave W, Montreal,  
QC, H3A 1A1, Canada  
Tel +1-514-934-1934 ext. 36922  
Fax +1-514-843-1493  
E-mail [nancy.mayo@mcgill.ca](mailto:nancy.mayo@mcgill.ca)

**Background and Purpose** There is a variety of stroke risk factors, and engaging individuals in reducing their own personal risk is hugely relevant and could be an optimal dissemination strategy. The aim of the present study was to estimate the stroke risk for specific combinations of health- and lifestyle-related factors, and to develop a personalized stroke-risk assessment tool for health professionals and the general population (called the MyRisk\_Stroke Calculator).

**Methods** This population-based, longitudinal study followed a historical cohort formed from the 1992 or 1998 Santé Québec Health Surveys with information for linkage to health administrative databases. Stroke risk factors were ascertained at the time of survey, and stroke was determined from hospitalizations and death records. Cox proportional hazards models were used, modeling time to stroke in relationship to all variables.

**Results** A total of 358 strokes occurred among a cohort of 17805 persons (men=8181) who were followed for approximately 11 years (i.e., ~200000 person-years). The following regression parameters were used to produce 10-year stroke-risk estimates and assign risk points: for age (1 point/year after age 20 years), male sex (3 points), low education (4 points), renal disease (8 points), diabetes (7 points), congestive heart failure (5 points), peripheral arterial disease (2 points), high blood pressure (2 points), ischemic heart disease (1 point), smoking (8 points), >7 alcoholic drinks per week (3 points), low physical activity (2 points), and indicators of anger (4 points), depression (4 points), and anxiety (3 points). According to MyRisk\_Stroke Calculator, a person with <50, 75, and 90 risk points has a 10-year stroke risk of <3%, 28%, and >75%, respectively.

**Conclusions** The MyRisk\_Stroke Calculator is a simple method of disseminating information to the general population about their stroke risk. **J Clin Neurol 2014;10:1-9**

**Key Words** stroke risk, risk calculator, dissemination, risk factors, personalized stroke risk reduction.

## Introduction

Prevention is the most effective way of reducing the effects of stroke for both individuals and society. There is considerable knowledge about stroke risk factors, including the lifestyle factors of unhealthy diet, obesity, physical inactivity,

smoking, and excessive alcohol consumption, and psychological distress factors that can lead to elevated blood pressure and glucose, and dyslipidemia.<sup>1</sup> Knowledge of the biological profile will guide the decision-making in clinical management. However, improving the health of a population requires individuals to be engaged in reducing their own personal risk. Thus, knowledge of risk factors directly under the control of the person is highly relevant; nevertheless, current studies indicate that the majority of the population does not have this knowledge.<sup>2</sup> Communicating risk to the general population requires simple tools, with relevance to the typical

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

individual who is likely to have more than one risk factor or be engaged in risky behavior.<sup>3</sup> The guidelines of the American Heart Association regarding primary prevention in stroke concur with this, and state that risk assessment is important for the general public.<sup>4</sup>

A structured search of the literature using PubMed was performed to evaluate current risk-factor profiles using the following keywords: probability, risk stroke, and behavioral. None of the identified risk profiles included physical activity, alcohol consumption, smoking, or psychological distress.<sup>3,5-9</sup> The best model is from the Framingham Study and is based on sex, age, systolic blood pressure, diabetes mellitus, cigarette smoking, prior cardiovascular disease, atrial fibrillation (AF), left ventricular hypertrophy, and use of antihypertensive medication.<sup>7</sup> Understanding these factors is not within the grasp of most people, and waiting until these conditions appear is not the optimal public health strategy. A useful stroke-risk profile based on lifestyle-related factors indicates approaches, such as increasing physical activity, that are within the patient's own control and would modify risk at an earlier stage.<sup>5,6,10</sup>

The aim of this study was to develop a prototype for a personalized stroke-risk assessment tool (called the MyRisk\_Stroke Calculator) based on risk estimates derived from the general population in Quebec, Canada. The specific objective of this research was to estimate the risk of stroke for specific combinations of health- and lifestyle-related variables, including age, sex, cardiovascular comorbidities, physical activity, alcohol consumption, smoking status, and psychological distress factors.

## Methods

### Participants and study design

This was a population-based, prospective, cohort study that used historical survey data from the 1992 and 1998 Santé Québec Health Surveys or "Enquête Sociale et de Santé",<sup>11</sup> which were linked to the hospital discharge database (MedEcho), the medical services databases maintained by the Régie de l'Assurance Maladie du Québec (RAMQ), and the death registry. The surveys were interviewer-assisted, and covered sociodemographics and lifestyle behaviors related to health and certain aspects of mental health.<sup>11</sup> Persons with an indication on the survey of previous stroke, a hospitalization record, or two or more stroke-related physician billings during the year prior to cohort entry were excluded. The remaining individuals were followed from the date of the survey (either 1992 or 1998) up to and including 2007.

The study received ethical approval from the McGill Faculty of Medicine Institutional Review Board (IRB Review Num-

ber A10-M76-03B), following approval from Quebec's Access to Information Bureau (Commission d'Accès à l'Information du Québec). Access to data was obtained only from persons who provided written consent to Santé Québec to link the survey data to health administration records.

### Measurement of stroke status

All persons with a hospitalization (in the MedEcho database) or death record indicating acute stroke based on International Classification of Disease version 9 (ICD-9) codes 431, 434, 436, and 362.3 were included.<sup>12,13</sup> These codes have been shown to accurately detect stroke in Quebec.<sup>12</sup> The corresponding ICD-10 codes are I61, I63, I66, I64, I67.6, and H34. Self-reports and data from RAMQ and MedEcho were used to identify people with an indication of past cerebrovascular disease using ICD-9 codes 430-438.

### Measurement of stroke risk factors

Information on health conditions present at the time of survey and during the previous year was available from three sources: questions from the survey, record of a physician billing (RAMQ), and hospitalization data (MedEcho). The following prevalent vascular conditions known to lead to an increased stroke risk were included<sup>4</sup>: AF, congestive heart failure, diabetes, hypertension, ischemic heart disease, peripheral arterial disease, and renal disease.

Body mass index (BMI) was available. Alcohol consumption was categorized into never, past, and amount consumed per week (1-6, 7-13, and 14 or more standard drinks). The frequency of weekly physical activity was counted as the number of sessions of 20-30 minutes of nonstop physical activity.<sup>11</sup> Subjects were categorized as smokers, which included regular (every day) and occasional (less than once a day) smoking, former smokers, and never smokers.<sup>11</sup>

Also included was the short version of Illfeld's Psychiatric Symptom Index,<sup>14</sup> validated in French,<sup>15</sup> which measures the presence of depressive or anger-related symptoms indicating the need for referral to a mental-health professional. Items were combined into a total score. However, since total scores are not useable for the purpose of risk assessment, single binary indicator variables were created for each question using the endorsement of "often" or "always" as a cutoff criterion.

### Statistical analyses

SAS version 9.2 software (SAS Institute, 100 SAS Campus Dr, Cary, NC, USA) was used for all statistical analyses. Each risk factor was dichotomized into two levels corresponding to whether or not the risk factor was present at cohort entry. Cluster analysis, using the proc FASTCLUS, was employed to identify the distribution of the number of risk factors per

person and to identify common risk-factor profiles.

Cox proportional-hazards regression, using Proc PHREG, was utilized to model the time to stroke during the follow-up period, which was up to and including 2007. The first approach was to fit a “development” model to the 1992 data and a “validation” model to the 1998 data, thereby taking advantage of the two waves of the survey. To test the similarity between the two cohorts, a final model was constructed with survey wave as a dummy variable. The overall performance of the model was tested using the Hosmer-Lemeshow method.<sup>16</sup> Risk was categorized into 10 bins, and the following parameters were calculated for each of those bins: average predicted risk, observed and predicted number of strokes, and observed and predicted number of individuals without a stroke.

The fit was also assessed by calculating the likelihood displacement (LD) statistic,<sup>17</sup> which is the amount by which minus twice the log (-2 log) likelihood, under a fitted model, changes when each subject is left out in turn. A critical value outside  $\pm 1.68$  suggests an influential subject.

All except three of the available variables were included in the primary model. The *a priori* excluded variables were AF, because the rarity of this diagnosis was such that no men in the present sample with AF had stroke; BMI, as the only significant association was if BMI was missing; and self-reported health, as there would be no public health action on this variable. The final model included those variables that, in the presence of all other variables, had a significant effect on stroke occurrence. Two other criteria were also used to avoid the bias of variable selection being sample-driven: 1) variables where the lower limit of the 95% confidence interval (CI) around the point estimate exceeded 0.90, or 2) the presence of strong literature evidence that the variable increases the risk of stroke.

For the calculation of individual risk, the regression parameters ( $\beta$ ) from the fully adjusted Cox model were used to derive weights for each risk factor according to the formula  $\beta \times 10$ , and rounded to the nearest whole digit, where hazard ratio (HR) =  $e^\beta$ . The weights were then summed to calculate the total risk score.

## Results

There were no significant differences in the risk estimates from the “development” (1992) and “validation” (1998) cohorts for any risk factor, indicating the reliability of the risk estimates. Hence, both cohorts were combined to create a total cohort for the final model. The calculated probability and observed probability of stroke were compared (Fig. 1). The overall chi-square value (df=7) from the Lemeshow and Hos-

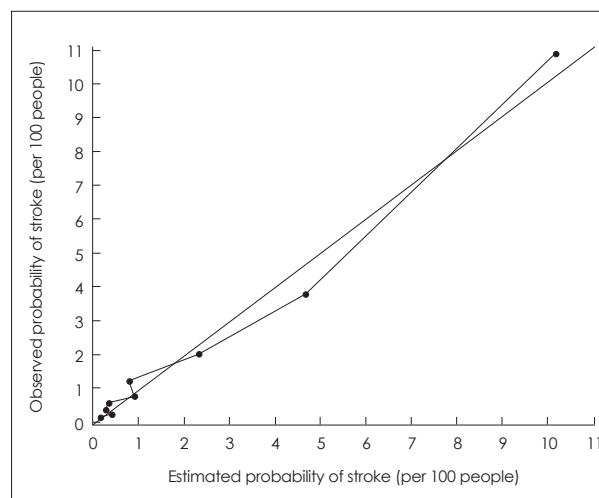


Fig. 1. Calculated and observed probabilities of stroke.

mer<sup>16</sup> test was 12.57, which was associated with a *p* value of 0.08 and hence provided no significant evidence that the model does not fit. The mean of the LD statistic<sup>17</sup> in the data set was 0.0014, with 99% of the subjects having an LD value of <0.05. Five people had an LD value between 0.24 and 0.35. These results do not indicate that there were any overly influential individuals affecting the model parameters.

Table 1 provides information on the total cohort. The estimated rates of stroke are compatible with rates for Quebec, as reported previously.<sup>18</sup> The proportions of the total population with more than one vascular comorbidity, more than one lifestyle factor, and more than one psychological distress indicator were 18%, 66%, and 20%, respectively (data not shown). Nearly 50% of those who developed stroke had no vascular comorbidity at the time of cohort entry.

Table 2 presents for each risk factor the prevalence, HR (fully adjusted models) and 95% CI, the parameter estimate, and the points allocated to each factor. For example, 45.9% of the population was male, and their stroke risk compared to women (adjusted for all variables in the model) was 1.33. The  $\beta$  used to derive the HR was 0.28, and multiplication by 10 and rounding indicated that the number of points to be used in the MyRisk\_Stroke Calculator was 3. The HR and  $\beta$  are presented for each year of age (after age 20 years); 1 point is assigned for every year older than 20.

Table 3 presents the format for entering risk information into the MyRisk\_Stroke Calculator, enabling a person to calculate their personal 10-year risk of stroke.

Table 4 is a look-up table relating cumulated points on the MyRisk\_Stroke Calculator to risk of stroke. The 10-year risk of stroke was calculated based on the Breslow baseline survival function, as recommended by Hanley,<sup>19,20</sup> which was found to be 0.99982. This corresponds to the probability that a 20-year old healthy woman (with no vascular comorbidities,

**Table 1.** Characteristics of the cohort

	Women		Men	
	Number (SD)	%	Number (SD)	%
Surveyed	23868	-	22606	-
Provided health insurance number	9732	40.8	8282	36.6
Previous stroke	108	-	101	-
Patient data analyzed	9624	-	8181	-
Average years of follow-up	11.3 (3.1)	-	11.1 (3.3)	-
Person-years	108948	-	91050	-
No. of strokes	175	-	183	-
Rate per 100000 person-years	160.6	-	201.0	-
Age (years)				
20–49	6239	64.8	5345	65.3
50–64	2116	22.0	1856	22.7
65–79	1136	11.8	907	11.1
80+	133	1.4	73	0.9
Total	9624	100	8181	100
Age at stroke (years)				
Mean	72.6 (11.7)	-	67.8 (12.9)	-
Median	73	-	70	-

no lifestyle or psychological risk indicators, and with a university degree) would survive without a stroke for 10 years. The risk of stroke was calculated using the following formula:

$$\text{Risk} = 1 - [S_{\text{reference}}(t)]^{\exp(\text{total score on MyRisk\_Stroke}/10)}$$

Therefore, an individual with a total score of 30 would have the following 10-year risk of stroke:

$$\text{Risk} = 1 - [0.99982]^{\exp(30/10)} = 0.00361 \text{ (or nearly 4 chances in 1000).}$$

## Discussion

The described process linking survey data to health records for a 10- to 14-year period has identified the risk factors for the occurrence of stroke and allowed the creation of a tool for an individual to quantify his or her stroke risk over the next 10 years based on a personal risk profile. A unique feature of the MyRisk\_Stroke Calculator is the inclusion of lifestyle factors, such as smoking, physical activity, and alcohol consumption, and psychosocial variables such as mood, anger, and anxiety, on top of well-established vascular comorbidities.

The risk predictions for the MyRisk\_Stroke Calculator were similar to those from the published literature for the effects of congestive heart failure, diabetes, hypertension, ischemic heart disease, peripheral arterial disease, and chronic kidney disorder.<sup>3,7,21–25</sup> Lifestyle risks were also similar to those reported in the literature.<sup>26–28</sup>

Indicator questions related to the constructs of depression,

anger, and anxiety were used to measure the effects of psychosocial variables. An individual is able to answer a directly posed question about a symptom or sign, but may have more difficulty providing a rating of the degree to which they have the latent constructs. Risk estimates for clinically diagnosed major depression from the literature are roughly 3.6<sup>29</sup> among individuals under the age of 65 years, and 3.2 among men.<sup>30</sup> The HR for the indicator variable “cry easily” was 1.44 (95% CI=0.98–2.09) in the present study. Similarly, the reported (adjusted) relative HR for anger expression and stroke is 2.0.<sup>31</sup> Furthermore, the estimate for the indicator variable “critical of others” was 1.45 (95% CI=0.95–2.22).

While some studies have found that obesity (as measured by BMI) increases the risk of stroke,<sup>5,6</sup> other studies have failed to show an association,<sup>32,33</sup> particularly after adjustment for factors such as the history of cardiovascular disease, diabetes mellitus, and hypertension.<sup>43</sup> The only category of BMI that was a significant predictor of stroke in the present study was “missing” (which may reflect a high BMI), and hence BMI was not included.

It is noteworthy that 50% of strokes occurred in individuals who had none of the vascular comorbidities in the MyRisk\_Stroke Calculator. This is similar to the results of Whisnant, who found that the combination of hypertension, transient ischemic attack, smoking status, ischemic heart disease, AF, diabetes, and mitral valve disease accounted for only 57% of ischemic strokes.<sup>34</sup> However, lifestyle factors are rarely included in risk profiles. Ezzati et al.<sup>35</sup> found that 70–76% of strokes could be attributed to the joint effects of hypertension, hypercholesterolemia, obesity, low fruit and

vegetable consumption, physical inactivity, and smoking. The INTERSTROKE study found that lifestyle, vascular conditions, psychosocial stress, depression, and the ratio of apolipoproteins B to A1 had a combined population attributable risk of 90.3%.<sup>36</sup> That study did not include education or

other measures of socioeconomic status. Although education is not necessarily a modifiable risk factor, excluding it could negatively affect the accuracy of prediction by underestimating the risk among the disadvantaged, potentially exacerbating health inequalities.<sup>37</sup> The predictions obtained in the pres-

**Table 2.** Parameter estimates for the variables included in the final Cox proportional-hazards model for stroke risk

Factor	Prevalence (%)	HR	95% confidence interval	$\beta$ estimate	Points used in the MyRisk_Stroke Calculator
Sex					
Women	54.05	0.75	0.59–0.95	0.28	3
Men	45.95		Reference		
Cohort 1992 (development)	58.64	1.13	0.89–1.45	0.13	
Cohort 1998 (validation)	41.36		Reference		
Age (per 1 year from age 20 years)	-	1.10	1.09–1.11	0.09	1
Congestive heart failure	0.94	1.70	1.02–2.82	0.53	5
Diabetes	3.64	1.95	1.43–2.67	0.67	7
Hypertension	12.52	1.18	0.93–1.50	0.16	2
Ischemic heart disease	5.43	1.08	0.80–1.46	0.08	1
Peripheral arterial disease	0.47	1.19	0.58–2.48	0.18	2
Renal disease	0.54	2.21	1.15–4.25	0.79	8
Smoking status					
Smoker	34.36	2.24	1.63–3.08	0.81	8
Former	34.88	1.59	1.18–2.14	0.46	5
Never	27.41		Reference		0
Missing	3.35	1.54	0.94–2.52	0.43	
Physical activity					
Inactive	47.69	1.20	0.92–1.55	0.18	2
Active 1 or 2 times per week	25.66	1.15	0.83–1.58	0.14	1
Active 3 or more times per week	24.88		Reference		0
Missing	1.77	1.08	0.61–1.92	0.08	
Alcohol consumption					
7 or more drinks per week	18.15	1.36	0.98–1.88	0.30	3
1–6 drinks per week	58.17		Reference		0
Former	6.58	1.69	1.22–2.34	0.53	5
Abstain	13.10	1.28	0.95–1.70	0.24	2
Missing	4.01	1.26	0.81–1.97	0.24	
Fearful					
Often or always	7.21	1.40	0.91–2.15	0.34	3
Sometimes, never, or missing	92.79		Reference		0
Cry easily					
Often or always	7.76	1.42	0.97–2.07	0.35	4
Sometimes, never, or missing	92.24		Reference		0
Critical of others					
Often or always	5.29	1.45	0.95–2.22	0.37	4
Sometimes, never, or missing	94.71		Reference		0
Education					
No high-school diploma	23.61	1.48	1.10–1.97	0.39	4
High-school diploma and/or some college	40.74	1.28	0.98–1.66	0.25	3
College or university or professional diploma	35.02		Reference		0
Missing	0.63	1.36	0.49–3.74	0.30	

**Table 3.** The MyRisk\_Stroke Calculator

	Your score
<b>Nonmodifiable factors</b>	
1. How old are you? For each year above the age of 20 you get 1 point (i.e., someone who is 45 would get 25×1=25 points).	<input type="text"/>
2. What is your gender? If you are male you get 3 points.	<input type="text"/>
3. What is the highest level of education you have completed? If you have completed some secondary schooling without a diploma you get 4 points. If you have a secondary diploma or some postsecondary education you get 3 points. If you have a postsecondary diploma you get 0 points.	<input type="text"/> <input type="text"/> <input type="text"/>
4. Have you ever been diagnosed with any type of renal or kidney disease? If yes you get 8 points, if no you get 0 points.	<input type="text"/>
5. Have you ever been diagnosed with diabetes? If yes you get 7 points, if no you get 0 points.	<input type="text"/>
6. Have you ever been diagnosed with congestive heart failure? If yes you get 5 points, if no you get 0 points.	<input type="text"/>
7. Have you ever been diagnosed with peripheral arterial disease? If yes you get 2 points, if no you get 0 points.	<input type="text"/>
8. Have you ever been diagnosed with or received medication for high blood pressure? If yes you get 2 points, if no you get 0 points.	<input type="text"/>
9. Have you ever been diagnosed with ischemic heart disease (e.g., a heart attack, myocardial infarction, or angina)? If yes you get 1 point, if no you get 0 points.	<input type="text"/>
<b>Modifiable factors</b>	
10. Which of the following best characterizes your smoking status? If you currently smoke you get 8 points. If you smoked in the past you get 5 points. If you have never smoked you get 0 points.	<input type="text"/> <input type="text"/> <input type="text"/>
11. Which of the following best characterizes your level of alcohol consumption? If you are a former drinker you get 5 points. If you drink more than 7 alcoholic beverages per week you get 3 points. If you have never had any alcohol you get 2 points. If you drink 1-6 drinks per week you get 0 points.	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
12. Which of the following best characterizes your current level of physical activity (where being physically active is defined as being active for at least 30 minutes)? If you have not been active this week you get 2 points. If you have only been active 1 or 2 times this week you get 1 point. If you have been active 3 or more times this week you get 0 points.	<input type="text"/> <input type="text"/> <input type="text"/>
13. How often do you feel critical of others? If you answered often or always you get 4 points. If you answered sometimes or rarely you get 0 points.	<input type="text"/>
14. How often do you cry easily? If you answered often or always you get 4 points. If you answered sometimes or rarely you get 0 points.	<input type="text"/>
15. How often do you feel fearful? If you answered often or always you get 3 points. If you answered sometimes or rarely you get 0 points.	<input type="text"/>
Your total score is:	<input type="text"/>

**Table 4.** Risk of stroke compared to score from the MyRisk\_Stroke Calculator

Total score (0–138)	10-year risk of stroke
0	0.00018
5	0.00030
10	0.00049
15	0.00081
20	0.00133
25	0.00219
30	0.00361
35	0.00594
40	0.00978
45	0.01607
50	0.02636
55	0.04309
60	0.07005
65	0.11285
70	0.17915
75	0.27782
80	0.41528
85	0.58718
90	0.76746
95	0.90973
100	0.98103
105	0.99998
110	1.00000

ent study, which included comorbidities, lifestyle, and stress-related factors in addition to education, was even stronger since when all factors were included in the MyRisk\_Stroke Calculator, only 1.7% of stroke cases were not explained.

The following scenario illustrates the application of the MyRisk\_Stroke Calculator to an individual. John is a 62-year-old man who is thinking forward to retirement. He did not finish postsecondary education, is a current smoker, regular drinker (corresponding to a heavy weekly alcohol consumption), and inactive. In addition, John has some anger issues. John was told that he was prediabetic during his annual physical assessment with his family doctor.<sup>38</sup> John's physician warned him that if he continues his present lifestyle choices he will develop full-fledged diabetes and his risk of heart disease or stroke will be high. John completes the MyRisk\_Stroke Calculator and receives 42 points due to age, 3 points for being male, 3 points due to his level of education, 8 points because of smoking, 3 points due to heavy alcohol consumption, 2 points due to inactivity, and 4 points because he expresses anger. Thus, John receives a total score of 65, which translates to an estimated 10-year risk of stroke of approximately 11.3% (Table 4). John wants to see what would happen to him if he becomes diabetic at the age of his retirement (65 years). In this future scenario, John's total score would be 75 (+7 for

diabetes and +3 for a projected age of 65 years), which gives him an estimated 10-year risk of stroke of 27.8%. With diabetes, his stroke risk at age 65 years is 2.5 times that without diabetes. John decides to follow the advice of his physician to reduce his risk of developing diabetes.

In order to motivate himself further, John looks to see how his risk of stroke would be lowered if, over the next 5 years until retirement, he became active (-2 points), reduced his drinking to four drinks per week (-3 points), and quit smoking (-3 points). With this new healthy lifestyle, his projected score at age 65 years is now 60 points, and his risk of stroke is reduced to 7.0%, which represents a substantial risk reduction from his projected course without any lifestyle change (27.8% with no changes to 7% with changes).

Using the Framingham Stroke Calculator<sup>7</sup> (assigning untreated systolic blood pressure as 126–135 mm Hg), the risks for these scenarios would be 8% for the status quo, 11% for developing diabetes by age 65 years, and 5% with a modified lifestyle and without diabetes. The risks predicted from the Framingham calculator are somewhat lower than those predicted using the MyRisk\_Stroke Calculator because of differences in the weights for smoking and diabetes, which possibly reflect inherent differences in the studied populations. In the Framingham stroke-risk calculator, diabetes and smoking are assigned 2 and 3 points, respectively; for the MyRisk\_Stroke Calculator, the corresponding values are 7 and 8 points, respectively. The MyRisk\_Stroke Calculator also includes alcohol consumption, inactivity, education, and psychological distress factors in the risk burden.

The above scenario illustrates several important novel features of the MyRisk\_Stroke Calculator, including the importance of considering both vascular comorbidities and lifestyle risk factors in the same calculator. Most individuals do not know their exact blood pressure or cholesterol levels, which are required for other stroke-risk calculators but not for the MyRisk\_Stroke Calculator, making it much more user-friendly. Importantly, by the time the individual presents with abnormalities in these factors, the damage has already been done to the cardiovascular system and may not be easily reversible.<sup>39</sup> In fact, in their guidelines for healthcare professionals from the American Heart Association and the American College of Cardiology, Grundy et al. recommend that “appropriate intervention, guided by risk assessment that is performed periodically in early adulthood and early middle age, has the potential to bring about a significant reduction in long-term risk.”<sup>39</sup> As younger individuals are unlikely to already have significant abnormalities on clinical markers of atherosclerotic disease, the MyRisk\_Stroke Calculator provides a relevant tool for them to assess future stroke risk and take preemptive action. Approaches including lifestyle

factors were based on fixed risk profiles,<sup>6,9</sup> and did not permit isolation of the contribution of each risk factor. In the aforementioned scenario, smoking contributes much more to John's increased risk of stroke than does his physical inactivity, which may help him prioritize lifestyle changes.

An important caveat for any risk calculator is that the risk of any condition is not necessarily elastic, and so causal inferences should not be made solely on the basis of the MyRisk\_Stroke Calculator.<sup>8</sup> This means that individuals who successfully modify several risk factors do not necessarily attain the risk of an individual who never had that condition or risk indicator.

Data on several established risk factors for cardiovascular disease, such as cholesterol, glucose, and blood pressure values and diet were not available in this study. However, Gaziano<sup>10</sup> demonstrated that the inclusion of cholesterol level in a prediction model for risk of stroke did not significantly affect the fit of that model among either men or women. Vascular comorbidities were identified using records from hospitalizations, physician visits, and self-reports in the survey at cohort entry, which may have resulted in some measurement error. The development of comorbidities over the follow-up period were also not included, since it was felt that these ensued from the baseline risk profile, and hence formed part of the causal pathway. They were also treated as binary variables, and the effect of severity was not taken into account. The low-risk weight assigned to some vascular comorbidities likely reflects treatment effectiveness. Occurrences of stroke were identified using records from the hospitalization database. Several previous studies, including some in Quebec, have validated the use of ICD-9 codes to identify strokes in many different populations, and found high positive predictive values, negative predictive values, and sensitivity.<sup>12,40-43</sup>

The measurement of stroke status was particularly important in this study, since all risk estimates are based on the number of strokes identified. If the number of strokes that occurred in the sample were substantially underestimated, then all of the risk estimates provided by the MyRisk\_Stroke Calculator would be underestimated. However, use of the hospital discharge and death-certificate databases allowed us to follow this large sample and ensured a very low dropout rate during the follow-up. The MyRisk\_Stroke Calculator met 13 of the 17 criteria of a clinical prediction rule defined by Laupacis et al.<sup>44</sup> Although the Lemeshow and Hosmer<sup>16</sup> method was used to evaluate how well the model was able to discriminate between individuals at a high and low risk of stroke, further studies are needed to prospectively validate the MyRisk\_Stroke Calculator in different populations. The cohort in the present study was relatively young and there were few events, thus possibly limiting the ability to include interaction effects

in this model due to an insufficient statistical power, which may limit the generalizability of the MyRisk\_Stroke Calculator to older individuals. Therefore, the MyRisk\_Stroke Calculator is not necessarily appropriate for use in individuals older than 75 years; a 5-year stroke risk may be more appropriate for that population. Although a split-sample approach involving "development" and "validation" cohorts was used to develop and validate the estimates for the MyRisk\_Stroke Calculator, this reduces the power and introduces instability into the estimates. Therefore, the estimates from the model were reported with both surveys combined. Before this prototype is ready for application for personal use or in clinical practice, it would need to be further validated, preferably in an independent sample.

The MyRisk\_Stroke Calculator is a simple-to-use method of disseminating risk information to the general population and may serve as an additional strategy for primary prevention. Physicians could use the calculator to open a dialogue with their patients and engage them actively in risk reduction. While prospective studies are needed to test the validity of the MyRisk\_Stroke Calculator, the engagement of people in reducing their own risk of stroke is an ongoing challenge and needs to be the focus of research into knowledge translation.

#### Conflicts of Interest

The authors have no financial conflicts of interest.

#### Acknowledgements

We thank the senior biostatistician Susan Scott for technical assistance in SAS programming.

#### REFERENCES

1. World Health Organization. *Preventing Chronic Diseases: A Vital Investment*. Geneva: World Health Organization, 2005.
2. Schneider AT, Pancioli AM, Khoury JC, Rademacher E, Tuchfarber A, Miller R, et al. Trends in community knowledge of the warning signs and risk factors for stroke. *JAMA* 2003;289:343-346.
3. Yusuf HR, Giles WH, Croft JB, Anda RF, Casper ML. Impact of multiple risk factor profiles on determining cardiovascular disease risk. *Prev Med* 1998;27:1-9.
4. Goldstein LB, Adams R, Alberts MJ, Appel LJ, Brass LM, Bushnell CD, et al. Primary prevention of ischemic stroke: a guideline from the American Heart Association/American Stroke Association Stroke Council: cosponsored by the Atherosclerotic Peripheral Vascular Disease Interdisciplinary Working Group; Cardiovascular Nursing Council; Clinical Cardiology Council; Nutrition, Physical Activity, and Metabolism Council; and the Quality of Care and Outcomes Research Interdisciplinary Working Group: the American Academy of Neurology affirms the value of this guideline. *Stroke* 2006;37:1583-1633.
5. Kurth T, Moore SC, Gaziano JM, Kase CS, Stampfer MJ, Berger K, et al. Healthy lifestyle and the risk of stroke in women. *Arch Intern Med* 2006;166:1403-1409.
6. Meng L, Maskarinec G, Lee J, Kolonel LN. Lifestyle factors and chronic diseases: application of a composite risk index. *Prev Med* 1999;29:296-304.
7. Wolf PA, D'Agostino RB, Belanger AJ, Kannel WB. Probability of



- stroke: a risk profile from the Framingham study. *Stroke* 1991;22:312-318.
8. Kokotailo RA, Hill MD. Coding of stroke and stroke risk factors using international classification of diseases, revisions 9 and 10. *Stroke* 2005;36:1776-1781.
  9. Weikert C, Berger K, Heidemann C, Bergmann MM, Hoffmann K, Klipstein-Grobusch K, et al. Joint effects of risk factors for stroke and transient ischemic attack in a German population: the EPIC Potsdam study. *J Neurol* 2007;254:315-321.
  10. Gaziano TA. Reducing the growing burden of cardiovascular disease in the developing world. *Health Aff (Millwood)* 2007;26:13-24.
  11. *Enquête Sociale et de Santé 1998*. 2nd ed. Québec City: Gouvernement du Québec, 2001.
  12. Mayo NE, Chockalingam A, Reeder BA, Phillips S. Surveillance for stroke in Canada. *Health Rep* 1994;6:62-72.
  13. Spolaore P, Brocco S, Fedeli U, Visentin C, Schievano E, Avossa F, et al. Measuring accuracy of discharge diagnoses for a region-wide surveillance of hospitalized strokes. *Stroke* 2005;36:1031-1034.
  14. Ilfeld FW Jr. Further validation of a psychiatric symptom index in a normal population. *Psychol Rep* 1976;39:1215-1228.
  15. Kovess V, Murphy HB, Fournier L, Tousignant M. Theoretical and practical problems in the measurement of the mental health of populations. Quebec health survey. *Acta Psychiatr Belg* 1986;86:588-593.
  16. Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982;115:92-106.
  17. Escobar LA, Meeker WQ Jr. Assessing influence in regression analysis with censored data. *Biometrics* 1992;48:507-528.
  18. Mayo NE, Nadeau L, Daskalopoulou SS, Côté R. The evolution of stroke in Quebec: a 15-year perspective. *Neurology* 2007;68:1122-1127.
  19. Hanley JA. The Breslow estimator of the nonparametric baseline survivor function in Cox's regression model: some heuristics. *Epidemiology* 2008;19:101-102.
  20. Julien M, Hanley JA. Profile-specific survival estimates: making reports of clinical trials more patient-relevant. *Clin Trials* 2008;5:107-115.
  21. Kannel WB, Wolf PA, Verter J. Manifestations of coronary disease predisposing to stroke. The Framingham study. *JAMA* 1983;250:2942-2946.
  22. Abbott RD, Donahue RP, MacMahon SW, Reed DM, Yano K. Diabetes and the risk of stroke. The Honolulu Heart Program. *JAMA* 1987;257:949-952.
  23. Manson JE, Colditz GA, Stampfer MJ, Willett WC, Krolewski AS, Rosner B, et al. A prospective study of maturity-onset diabetes mellitus and risk of coronary heart disease and stroke in women. *Arch Intern Med* 1991;151:1141-1147.
  24. Leng GC, Lee AJ, Fowkes FG, Whiteman M, Dunbar J, Housley E, et al. Incidence, natural history and cardiovascular events in symptomatic and asymptomatic peripheral arterial disease in the general population. *Int J Epidemiol* 1996;25:1172-1181.
  25. Ninomiya T, Kiyohara Y, Kubo M, Tanizaki Y, Doi Y, Okubo K, et al. Chronic kidney disease and cardiovascular disease in a general Japanese population: the Hisayama study. *Kidney Int* 2005;68:228-236.
  26. Reynolds K, Lewis B, Nolen JD, Kinney GL, Sathya B, He J. Alcohol consumption and risk of stroke: a meta-analysis. *JAMA* 2003;289:579-588.
  27. Reimers CD, Knapp G, Reimers AK. Exercise as stroke prophylaxis. *Dtsch Arztebl Int* 2009;106:715-721.
  28. Shinton R, Beevers G. Meta-analysis of relation between cigarette smoking and stroke. *BMJ* 1989;298:789-794.
  29. Salaycik KJ, Kelly-Hayes M, Beiser A, Nguyen AH, Brady SM, Kase CS, et al. Depressive symptoms and risk of stroke: the Framingham study. *Stroke* 2007;38:16-21.
  30. Bos MJ, Lindén T, Koudstaal PJ, Hofman A, Skoog I, Breteler MM, et al. Depressive symptoms and risk of stroke: the Rotterdam study. *J Neurol Neurosurg Psychiatry* 2008;79:997-1001.
  31. Everson SA, Kaplan GA, Goldberg DE, Lakka TA, Sivenius J, Salonen JT. Anger expression and incident stroke: prospective evidence from the Kuopio ischemic heart disease study. *Stroke* 1999;30:523-528.
  32. Walker SP, Rimm EB, Ascherio A, Kawachi I, Stampfer MJ, Willett WC. Body size and fat distribution as predictors of stroke among US men. *Am J Epidemiol* 1996;144:1143-1150.
  33. Curb JD, Marcus EB. Body fat, coronary heart disease, and stroke in Japanese men. *Am J Clin Nutr* 1991;53:1612S-1615S.
  34. Whisnant JP. Modeling of risk factors for ischemic stroke. The Willis Lecture. *Stroke* 1997;28:1840-1844.
  35. Ezzati M, Hoom SV, Rodgers A, Lopez AD, Mathers CD, Murray CJ, et al. Estimates of global and regional potential health gains from reducing multiple major risk factors. *Lancet* 2003;362:271-280.
  36. O'Donnell MJ, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, et al. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet* 2010;376:112-123.
  37. Lang T. Ignoring social factors in clinical decision rules: a contribution to health inequalities? *Eur J Public Health* 2005;15:441.
  38. Genuth S, Alberti KG, Bennett P, Buse J, Defronzo R, Kahn R, et al. Follow-up report on the diagnosis of diabetes mellitus. *Diabetes Care* 2003;26:3160-3167.
  39. Grundy SM, Pasternak R, Greenland P, Smith S Jr, Fuster V. Assessment of cardiovascular risk by use of multiple-risk-factor assessment equations: a statement for healthcare professionals from the American Heart Association and the American College of Cardiology. *Circulation* 1999;100:1481-1492.
  40. Roumie CL, Mitchel E, Gideon PS, Varas-Lorenzo C, Castellsague J, Griffin MR. Validation of ICD-9 codes with a high positive predictive value for incident strokes resulting in hospitalization using Medicaid health data. *Pharmacoepidemiol Drug Saf* 2008;17:20-26.
  41. Rinaldi R, Vignatelli L, Galeotti M, Azzimondi G, de Carolis P. Accuracy of ICD-9 codes in identifying ischemic stroke in the General Hospital of Lugo di Romagna (Italy). *Neurol Sci* 2003;24:65-69.
  42. Leone MA, Capponi A, Varrasi C, Tarletti R, Monaco F. Accuracy of the ICD-9 codes for identifying TIA and stroke in an Italian automated database. *Neurol Sci* 2004;25:281-288.
  43. Tolonen H, Salomaa V, Torppa J, Sivenius J, Immonen-Räihä P, Lehtonen A, et al. The validation of the Finnish Hospital Discharge Register and Causes of Death Register data on stroke diagnoses. *Eur J Cardiovasc Prev Rehabil* 2007;14:380-385.
  44. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA* 1997;277:488-494.