Chronic Diseases in Canada

Volume 26, Number 1, Winter 2005

Public Health Agency of Canada
Agence de santé publique du Canada

Canada

Published by authority of the Minister of Health

ISSN 0228-8699

Aussi disponible en français sous le titre *Maladies chroniques au Canada*

Chronic Diseases in Canada

a publication of the Public Health Agency of Canada

Apologies

We extend our sincere apologies to Dr. Claire Infante-Rivard, whose name was inadvertently omitted from the masthead of Volume 25, No. 3/4 of *Maladies chroniques au Canada and from the Chronic Diseases in Canada* Web site. Both of these oversights have since been corrected.

Chronic Diseases in Canada (CDIC) is a quarterly scientific journal focussing on current evidence rele- vant to the control and prevention of chronic (i.e., non- communicable) diseases and injuries in Canada. The journal publishes a unique blend of peer-reviewed feature articles by authors from the public and private sectors that may include research from such fields as epidemiology, public/community health, biostatistics, behavioural sciences and health services. Authors retain responsibility for the content of their articles; the opinions expressed are not necessarily those of the CDIC Editorial Committee or of the Public Health Agency of Canada.

Subscription is free upon request.

When notifying us of a change of address, please enclose your old address label.

Chronic Diseases in Canada
Public Health Agency of Canada
130 Colonnade Road
Address Locator: 6501G
Ottawa, Ontario K1A 0K9

Fax: (613) 941-3605
E-mail: cdic-mcc@phac-aspc.gc.ca

Indexed in Index Medicus/MEDLINE, PAIS (Public Affairs Information Service) and EMBASE, the Excerpta Medica database.

This publication is also available online at http://www.phac-aspc.gc.ca/publicat/cdic-mcc

Influential observations in weighted analyses: Examples from the National Longitudinal Survey of Children And Youth (NLSCY)

Jennifer J Macnab, JJ Koval, KN Speechley, and MK Campbell

Abstract

This paper highlights the impact of survey weights on model fit in multiple linear regression with specific reference to the National Longitudinal Survey of Children and Youth (NLSCY) and provides recommendations for the treatment of influential observations. Multiple linear regression was used to estimate the association between child and family factors in the preschool years and vocabulary development at school age. Analyses were performed with and without survey weights. The model fit was assessed by examining the distribution of the studentized residuals and the change in the regression coefficients that would occur if an observation were removed. Two summary measures of influence, Dffits and Cook's D are reported. The models were refit excluding influential observations. Weighting of the linear model resulted in previously non-influential observations having an undue influence on the estimation of the regression parameters in the weighted model. The influential observations were driven primarily by the size of the survey weight as opposed to unusual values of x and y. Researchers working with large national health surveys such as the NLSCY and the National Population Health Survey (NPHS) are advised to include a detailed influence analysis before any final conclusions are made.

Key words: *epidemiologic methods; health surveys; influential observations; linear models; regression analysis*

Introduction

Any survey that puts restrictions on sampling beyond those of simple random sampling is complex in design and requires special analytical considerations. There are two main issues in the analysis of complex surveys: 1) the use of sample weights to correct for differential representation and 2) the effect of sample design on estimates of sampling error. This paper focusses on the use of survey weights for valid parameter estimation and the potential implications for analysis.

Survey weights are used to reflect the differing probabilities of selection within each stratum. The principle behind estimation in a probability sample is that each person represents several other people. Each record is therefore weighted by the inverse of the probability of selecting the person. In addition, weights can be used to correct for differential response rates within subsamples and post-stratification can be used to adjust the sample distribution to a known population distribution. Ignoring the weights in the analysis of complex survey data will result in biased parameter estimates.

Some authors disagree about the necessity of weighted analysis. Because extreme weights in a relatively small number of clusters can greatly increase variance, Korn and Graubard[1] suggest assessing the impact of weighting on model efficiency. If weighted analysis results in an unacceptable increase in variance, then an unweighted analysis incorporating design variables may be more appropriate. This approach is only possible if an appropriate set of design variables is available for inclusion in the model. Stratum and cluster indicators are not routinely available from Statistics Canada or through the Remote Data Centres (RDCs). One of the reasons for this is that the sampling strategy is so complex that it cannot be easily modelled with a discrete set of design variables, and therefore the analyst wishing to take advantage of rich data sources such as the National Longitudinal Survey of Children and Youth[2] (NLSCY) or the National Population Health Survey (NPHS) must use a model that incorporates the survey weights provided by Statistics Canada.

In the literature, most of the discussion on weighted analysis has been focussed on the impact of weighting on model efficiency. Marked differences in the parameter estimates between weighted and unweighted analyses are assumed to be due to important missing covariates. Including these covariates in the model should minimize the difference between weighted and unweighted parameter estimates. There has been limited discussion, however, of the impact of survey

Author References

Jennifer J Macnab, Department of Epidemiology and Biostatistics, Faculty of Medicine and Dentistry, University of Western Ontario, London, Ontario, Canada
JJ Koval, KN Speechley, MK Campbell, Department of Epidemiology and Biostatistics, Faculty of Medicine and Dentistry, University of Western Ontario, London, Ontario, Canada
Correspondence: Jennifer J Macnab, Department of Epidemiology and Biostatistics, Faculty of Medicine and Dentistry, Kresge Building, Room K201, University of Western Ontario, London, Ontario, Canada N6A 5C1; Fax: (519) 850-4222; E-mail: jmacnab@uwo.ca.

weights on the potential influence that a single observation may have on the model fit.

In the analysis of any data set there tend to be a few observations that lie at the extremes of the distribution of the variable(s) of interest. In a multiple linear regression, it is the distribution of the residuals that is of primary interest. Large residuals, although indicative of poor fit, are acceptable as long as they do not have an undue influence on the estimate of the parameters. If the outlying observation has an *x*-value that is at the centre of the X distribution, it will have little impact on the least-squares fit. If however, the outlying observation has an *x*-value that is at the extreme of the X distribution, it will exert strong leverage on the regression coefficients because it is more important in determining the model fit. The combination of high leverage (at the extreme of the X-distribution) with a regression outlier can produce substantial influence on the estimate of the regression coefficients.[3,4] In a large data set, a detailed analysis of influential observations is often unnecessary because the large number of observations means that no one observation has the potential to have an undue influence on the least-squares fit. However, in the analysis of survey data, the degree of influence that an observation may have on the estimate of the parameters is determined not only by the size of the residual and the leverage of the *x*-value, but also by the size of the survey weight assigned to the outlying observation. Individuals who, as a result of sample design, have been assigned a large weight have a greater potential to influence the least-squares fit than individuals who represent only a very small proportion of the population.

Analysis of survey data from Statistics Canada is commonly a two-step procedure. The preliminary model building is done using a weighted *model-based* analysis, such as can be achieved using SAS PROC REG with a WEIGHT statement. The parameter estimates (*b*) from this model are valid but the standard errors are an underestimate of the true variance. Once the final model has been correctly specified, a *design-based* bootstrap variance estimation procedure can be used to obtain the correct standard errors. Example SAS and SPSS code for bootstrap variance estimation procedures are available from Statistics Canada. This paper describes procedures for the preliminary model-based analysis and demonstrates how a comparison of weighted and unweighted models can be a useful tool during preliminary analysis for identifying problems in model fit. The impact of weighting was examined by comparing the regression parameter estimates and regression diagnostics for both models. Influential observations were identified by a careful evaluation of the sample distribution of the studentized residuals (*Rstudent*) and the change in the estimated regression coefficients that would occur if an observation were removed.

Methods

Data Source

The data for this study came from the master data files for cycle 1 and cycle 2 of the National Longitudinal Survey of Children and Youth[2] (NLSCY) housed at Statistics Canada in Ottawa, Ontario. The NLSCY is a joint project of Human Resources Development Canada and Statistics Canada. It is a population-based longitudinal survey designed to monitor the development and well-being of Canada's children as they grow from infancy to adulthood. The target population for cycle 1 was Canadian children from newborn to 11 years of age who lived

TABLE 1
Unweighted and weighted means and proportions for the child and family predictors in boys (n = 1,115)

	Unweighted	Weighted	
	Mean/ Proportion	Mean/ Proportion	SE[b]
Outcome			
PPVT-R	99.78	98.10	0.61
Child predictors			
SGA (< 10 percentile)	0.06	0.05	0.01
LGA (≥ 90 percentile)	0.14	0.14	0.01
Premature (< 259 days)	0.10	0.11	0.09
Fussiness	16.70	16.70	0.36
Persistence	26.53	26.48	0.41
Externalizing behaviour	10.89	10.45	0.25
Internalizing behaviour	4.58	4.55	0.15
Motor social development	97.98	96.62	0.69
Good health	0.87	0.88	0.01
Family predictors			
Single parent status	0.12	0.19	0.02
Family size ≥ 4	0.06	0.04	0.01
PMK[a] Age³ 25	0.89	0.90	0.01
Education < High school	0.13	0.17	0.02
College/University[a]	0.43	0.40	0.02
Low income (< LICO)	0.21	0.26	0.02
Social support (< 10 percentile)	0.06	0.09	0.02
Family dysfunction	7.63	7.67	0.23
Covariates			
Small city (< 100,000)	0.31	0.21	0.02
Rural area	0.29	0.21	0.02
Foreign born	0.08	0.10	0.02
New Immigrant (< 10 yrs)	0.03	0.04	0.01

[a] PMK = Person Most Knowledgeable. In 91.3% of cases this was the mother

[b] The standard errors for the weighted means were calculated using bootstrapped variance estimation procedures

in private households. The first cross-sectional sample was collected in 1994. Since that time subsequent cycles have been repeated every two years.

The calculation of the survey weights for the NLSCY was a multi-stage procedure. The main steps in the calculation of the survey weights are summarized here; for more complete details, the interested reader should refer to the NLSCY documentation.[2] The NLSCY used information collected from the Labour Force Survey[5] (LFS) to identify households with children under the age of 12. Therefore, the survey weights in the NLSCY incorporate the sampling strategy of the LFS. The weight for each record began with the basic weight (BW) from the LFS. This weight was the inverse of the probability of the dwelling being selected. Simply put, if 10 out of every 1000 households were sampled, the basic weight would be 100 (1000/10) in that each household in the sample represents 100 households in the population. The basic weight was then adusted for sub-sampling (SS), non-response (NR), and rural/urban (RU) distribution. If the number of dwellings in a cluster increased to the extent that it created field collection problems, then a smaller sample (sub-sample) was used to keep interviewer assignments manageable. The sub-sampling correction was equal to the ratio of the number of dwellings in the original cluster to the number of dwellings interviewed. There was a non-response rate of approximately 5% in the LFS. Survey weights were corrected for non-response by proportionately increasing the weights of responding households by the ratio of the number sampled to the number responding. Post-stratification was also used to correct the estimate of rural/urban counts to estimates derived from the census. The LFS sub-weight was therefore equal to the BW*SS*NR*RU. The LFS sub-weight was then adjusted for sampling features of the NLSCY. This included corrections for the number of rotations used from the LFS, updating of the sampling frame to adjust for changes within the three month lag between the LFS and the NLSCY, households with more than one family, households with more than four children, non-response, and post-stratification to the 1995 census estimates. The final weights provided with the NLSCY are at the level of the individual. Both cross-sectional and longitudinal weights are provided with each record. The weights provided with each record are expsansion weights in that they sum to the population size, not to the sample size.

Selection of Study Sample

This study was restricted to the analysis of longitudinal data from cycle 1 and cycle 2 for the 2- to 3-year-old age cohort. Predictor

TABLE 2
Linear regression parameter estimates for the preliminary unweighted and weighted model-based analysis of receptive vocabulary on child and family factors in boys

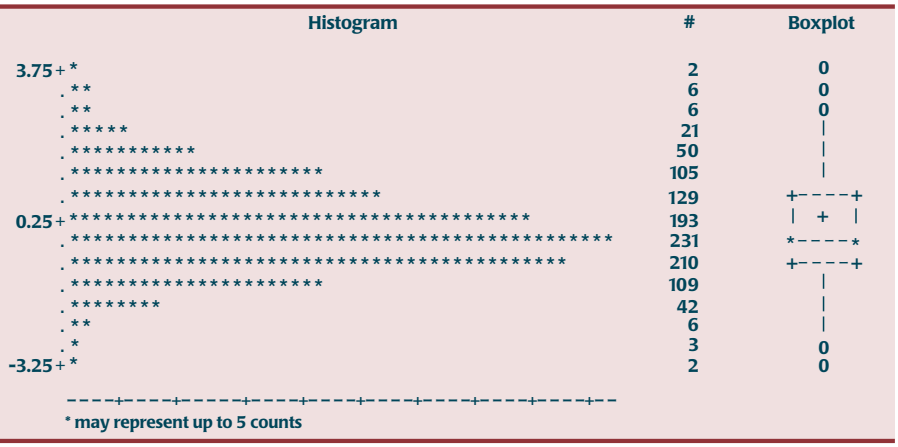
	Boys (n=1,115)		
	Unweighted	Weighted	
	b	*b*	% change
Intercept	81.80	73.21	
Covariates[a]			
Urban/rural area			
• Small city (< 100,000)	-1.49	-1.43	4.0
• Rural area	-2.27	-3.08	30.0
PMK[b] born outside of Canada/USA	**-1.74**	**-4.11**	**136.2**
New immigrant (< 10 yr)	-9.18	-6.92	24.6
Child			
Size for gestational age			
• SGA (< 10 percentile)	**-2.07**	**0.86**	**141.6**
• LGA (≥ 90 percentile)	-0.30	-0.35	16.7
Premature (< 259 days gestation)	-2.07	-2.04	1.5
Fussiness scale	**0.17**	**0.31**	**82.4**
Persistence scale	**-0.13**	**-0.03**	**76.9**
Externalizing behaviour scale	-0.24	-0.38	58.3
Internalizing behaviour scale	0.13	0.28	115.4
Motor social development scale	0.22	0.23	4.6
Good health	-0.91	0.60	165.9
Family			
Single parent status	**0.85**	**3.09**	**263.5**
Family size ≥ 4	-3.33	-3.36	0.9
PMK[b] age ≥ 25	3.71	5.71	53.9
PMK[b] education			
• < High school	**-2.59**	**-0.50**	**80.7**
• College/university	1.58	2.03	28.5
Income (< LICO)	**-3.76**	**-7.26**	**93.1**
Lack of social support (<10 percentile)	**-2.81**	**-1.90**	**32.8**
Family dysfunction	**-0.23**	**-0.37**	**60.9**

Note: Parameter estimates with more than a 75% change in the size of the coefficient between the unweighted and weighted models are bolded for emphasis.

[a] All models also controlled for province of residence coded as 9 dummy variables with Ontario as the reference level.

[b] PMK = Person Most Knowledgeable. In 91.3% of cases this was the mother

FIGURE 1
Distribution of studentized residual in unweighted male model of PPVT-R on child and family predictors



variables were collected from cycle 1 when the child was 2 to 3 years of age and outcome data were collected from cycle 2 when the child was 4 to 5 years of age. The study sample was restricted to singleton births and limited to one child per household. The final study sample consisted of 2,538 children who were age 2 to 3 years in cycle 1 (1994), and 4 to 5 years in cycle 2 (1996).

Analysis

All analyses were conducted by remote submission to Statistics Canada and executed on the master data file. To illustrate the impact of survey weights on parameter estimates in regression analysis, two preliminary multiple linear regressions were developed for the prediction of vocabulary development at age 4 to 5 years, as measured by the Peabody Picture Vocabulary Test[6] (PPVT-R), by child and family factors at age 2 to 3 years. The PPVT-R is a standardized test of receptive vocabulary with a mean of 100 and a standard deviation of 15. All analyses were conducted using SAS ver. 8.0. The first model was an unweighted linear regression using PROC REG. The second model included a WEIGHT statement that specified the longitudinal weight variable from the NLSCY data file. Standard errors for the final model were calculated using bootstrap variance estimation.

Differences in the parameter estimates between the unweighted and weighted model are expected because the unweighted model does not account for the unequal probability of selection within the sample. However, when the association under examination is not related to the sample design, the difference between the parameter estimates from each model should be minimal. Large differences, not easily explained by the sample design, should alert the analyst to potential problems in model fit during preliminary model building. In particular, the potential impact that survey weights can have on influential observations.

Unexpected differences between weighted and unweighted models can sometimes be accounted for by important missing covariates. Therefore, the model building began with comprehensive sets of child and family variables. The difference between the unweighted and weighted models can be even more pronounced if the missing covariates are strongly related to the sample design. For this reason, the models also included a set of geographical (province and rural/urban status) and socio-cultural covariates (parent's country of birth and immigration status). The inclusion, and subsequent elimination, of interaction terms tested for potential effect modification by province or rural/and urban area. Conceptually related variables were entered in blocks and the model was reduced (α = .20) to find the best subset of variables to represent the construct. The independent child variables selected were neonatal status (prematurity, size for gestational age), temperament (fussiness, persistence) behaviour (externalizing, internalizing) development at age 2 to 3 years (motor social development scale) and health status at age 2 to 3 years. The independent family variables included two measures of family structure (single parent status and family size) four measures of family resources (parental age, education, income, and social support) and one measure of family functioning.

Large differences between unweighted and weighted parameter estimates can also occur if the weighted model is misspecified. Therefore, residual scatterplots and partial residual plots were examined for evidence of non-linearity.[3] If indicated, higher order terms were entered and tested using

FIGURE 2
Distribution of dffits (Standard influence on predicted value) in unweighted male model of PPVT-R on child and family predictors

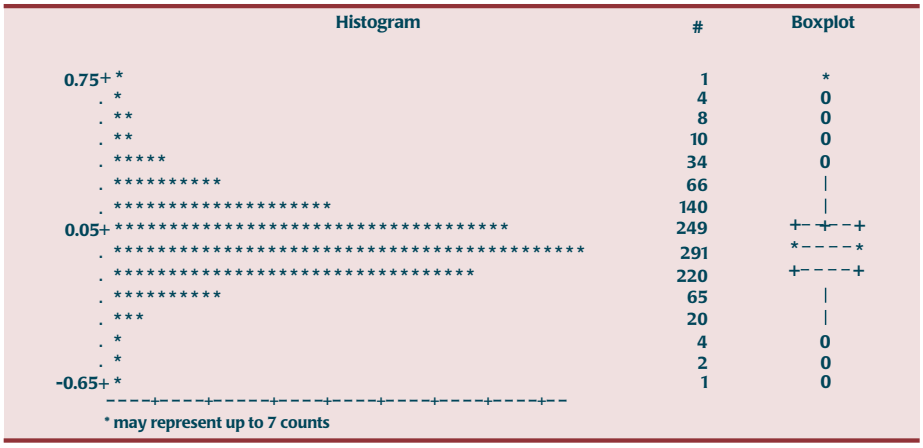


TABLE 3
Distribution of regression diagnostic statistics for both weighted and unweighted linear regression of competence on child and family predictors in boys: Starting model (n = 1,115) and model excluding influential observations (n = 1,100)

	Unweighted		Weighted	
	n = 1,115	n = 1,100[a]	n = 1,115	n = 1,100[a]
RSTUDENT[b]				
Mean	0.000	-0.020	0.007	-0.019
Median	--0.094	-0.070	0.007	-0.070
Min	-3.115	-3.978	-4.545	-3.978
Max	3.816	4.213	7.503	4.213
Dffits				
Mean	0.000	-0.009	-0.001	-0.009
Median	0.000	-0.008	-0.008	-0.008
Min	-0.610	-0.966	-2.531	-0.966
Max	0.794	0.920	2.978	0.920
Cooks D				
Mean	0.001	0.001	0.002	0.001
Median	0.001	0.000	0.000	0.000
Min	0.000	0.000	0.000	0.000
Max	0.020	0.030	0.272	0.030

[a] Excluding influential observations
[b] RSTUDENT = Studentized residual

backward elimination (α = .05). The correlation matrix variance inflation factors (*VIFs*) and condition indices (*CIs*) were examined for evidence of multicollinearity. Finally, effect modification by gender was tested by inclusion of interaction terms (predictor X gender). Five interaction terms remained in the model after backward elimination (α = .05), therefore the model was split on gender. For purposes of example, only the results for boys are reported.

The final step in the preliminary model building involved the identification of influential observations. The most direct measure of influence is the *Dfbeta,* which express the impact on each coefficient in the model of deleting each observation in turn in terms of a standardized coefficient. The problem with the examining *Dfbetas* in a large data set is their sheer number. For each model there are n(k + 1) *Dfbetas* values generated; where n = the sample size and k = the number of variables in the model. Because of the large number of *Dfbetas* generated for each model, two summary measures of fit were examined first to gain insight into the effect of deleting a single observation. If the summary measures indicated a single observation was influential then the nature of the influence was examined by looking at the *Dfbetas* for that observation.

Two summary measures of influence were examined: *Dffits* and *Cook's D*. *Dffits* provide a summary measure of the standardized change in all of the parameter estimates when a single observation is removed. *Cook's D* is a scale invariant distance measure of influence (discrepancy times leverage) on all the parameter estimates in the fitted model.[3] Criteria for defining influential values can be based on external scaling, internal scaling, and/or gaps in the distribution of the fit statistic.[7] External scaling provides cutoff values based on statistical theory. Standardized fit statistics, such as *Rstudent, Dffits,* and *Dfbetas,* are interpreted in the same way as z-scores and are considered large if their value exceeds 2. However, absolute cutoff values are based on assumptions about the distribution of the residuals and may risk missing relatively influential data. Internal scaling determines cutoff values based on an examination of the sample distribution of the fit statistics. Values that exceed 3.5 times the interquartile range are considered large.[7] Noticeable gaps in the sample distribution of the fit statistic

FIGURE3
Distribution of studentized residual in weighted male model of PPVT-R on child and family predictors

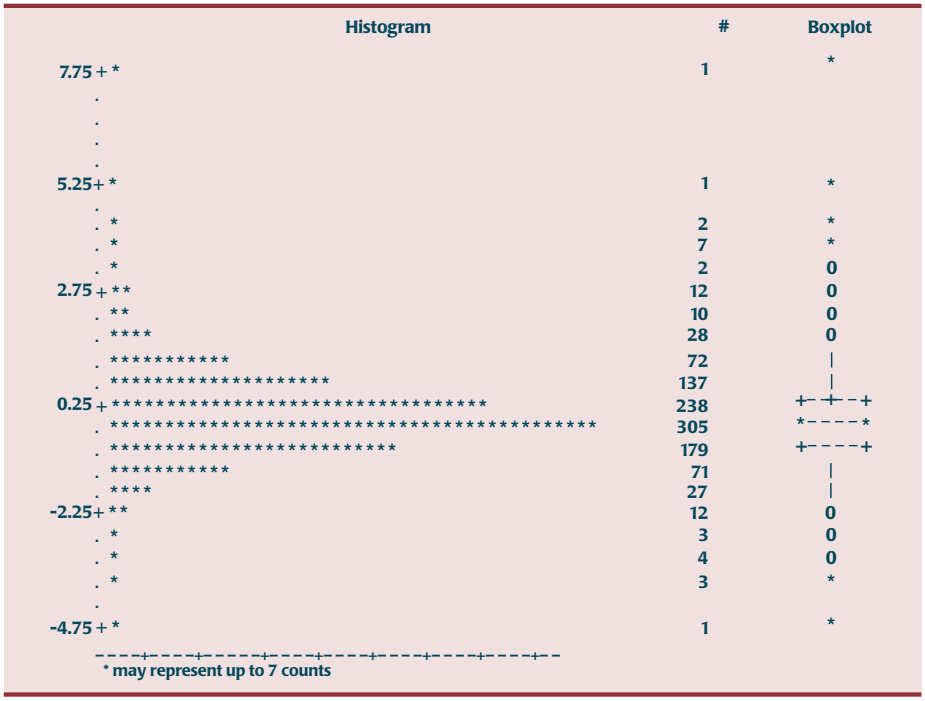


TABLE 4
Comparison of parameter estimates in unweighted and weighted linear regression models of competence on child and family variables with and without influential observations in boys.

	Unweighted		Weighted		
	n = 1,115	n = 1,100[a]	n = 1,115	n = 1,100[a]	S.E.[c]
Intercept	81.80	80.82	73.21	72.77	4.87
Covariates[b]					
Small urban (< 100,000)	-1.49	-1.02	-1.43	-0.72	1.18
Rural area	-2.27	-1.66	**-3.08**	**-1.68**	**1.35**
PMK birth	-1.74	-1.85	-4.11	-3.01	2.36
New immigrant (< 10 yr)	-9.18	-10.63	**-6.92**	**-10.75**	**3.44**
Child					
SGA < 10 percentile	-2.07	-2.64	**0.86**	**-2.10**	**2.55**
LGA ≥ 90 percentile	-0.30	-0.14	-0.35	-0.57	1.34
Premature (< 259 days)	-2.07	-2.30	-2.04	-3.33	1.76
Fussiness	0.17	0.16	0.31	0.23	.10
Persistence	-0.13	-0.13	-0.03	-0.04	.08
Externalizing behaviour	-0.24	-0.24	-0.38	-0.30	.12
Internalizing behaviour	0.13	0.18	0.28	0.34	.19
Motor social development	0.22	0.23	0.23	0.25	.03
Good health	-0.91	-0.93	0.60	0.57	1.61
Family					
Single parent status	0.85	-0.07	**3.09**	**0.67**	**1.92**
Family size ≥ 4	-3.33	-4.65	-3.36	-3.61	2.23
PMK age ≥ 25	3.71	3.43	**5.71**	**3.56**	**1.70**
Education < High school	-2.59	-3.08	**-0.50**	**-3.41**	**1.70**
Education > Hiigh school	1.58	1.89	**2.03**	**3.03**	**1.21**
Low income (< LICO)	-3.76	-3.20	**-7.26**	**-4.80**	**1.56**
Social support (< 10 percentile)	-2.81	-3.03	**-1.90**	**-3.62**	**1.57**
Family dysfunction	-0.23	-0.20	**-0.37**	**-0.22**	**.10**
Rsquare	0.19	0.20	0.23	0.27	
Adj R-Sq	0.16	0.18	0.22	0.25	

Note: Bolded parameter estimates indicate variables for which removal of the influential observations resulted in a change in the regression coefficient of greater than 1 standard deviation (s.e.).

[a] Excluding influential observations

[b] Model also controlled for province of residence.

[c] Standard error calculated using bootstrap variance estimation

can identify observations that are influential relative to the rest of the sample. Graphical analysis of the *Dffits* and *Cook's D* values were conducted to identify extreme values relative to the sample and gaps in the distribution. Based on the graphical analysis of the fit statistics, a standard criterion for identifying influential observations was developed. Observations with $|Dffits| > 1$ or *Cook's D* > .3 were defined as influential. Influential observations from the weighted analysis were removed and both models were repeated. The impact of removing the influential observations was examined by a comparison of the regression parameters in both the weighted and unweighted models.

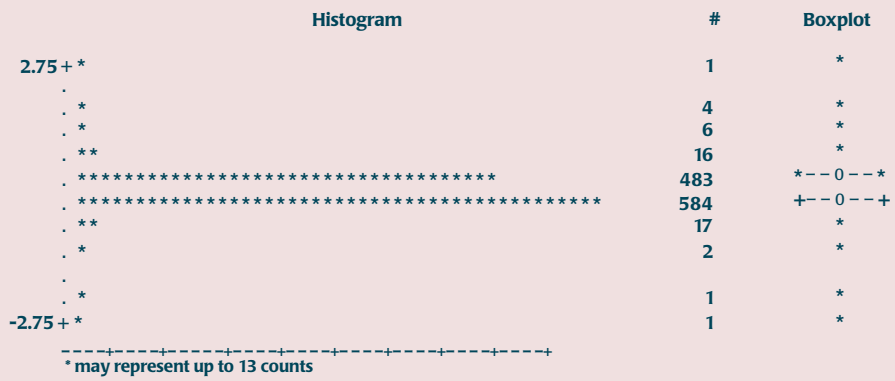
Results

Descriptive statistics for the variables selected for the regression analysis are shown in Table 1. There are only slight differences between the unweighted and weighted means and proportions for the child and family predictors. However, there is a substantial difference between the unweighted and weighted estimates for the geographical and socio-cultural covariates because they are closely associated with the probability of selection. The weights are largest for large cities within large provinces. When the weights are positively correlated with *x*, then the unweighted mean of *x* will be an underestimate of the population mean. Because single parent status, education less than high school, low income, and lack of social support are more common in large urban centres in large provinces, the unweighted proportions underestimate the true population proportions. Small urban centres and rural areas are over represented in the sampling frame and as such the unweighted estimates are overestimates of the true population parameters.

Table 2 shows the parameter estimates for the unweighted and weighted regression of vocabulary (PPVT-R) on child and family factors. There was a substantial difference (greater than a 75% change in the size of the regression coefficients) in the parameter estimates for small for gestational age (SGA), fussiness, persistence, internalizing behaviour, general health, single parent status, education, and income. Because of the marked difference between the parameter estimates from the unweighted and weighted models, both models were examined to determine if any observations were having an undue influence on the parameter estimates.

Figure 1 shows the distribution of the studentized residuals for the unweighted male model of vocabulary (PPVT-R) on child and family predictors. Although some residuals exceed 2, the distribution is relatively normal and as can be seen from Figure 2, none of the observations has an undue influence on the parameter estimates, in that the *dffit* values do not exceed .75. In the weighted model (see Figure 3), the studentized residuals ranged from –4.75 to 7.75 and the distribution was skewed to the right. The two observations with the largest residuals were previously not extreme values. The large residuals in the weighted model

FIGURE 4
Distribution of dffits (Standard influence on predicted value) in weighted male model of PPVT-R on child and family predictors



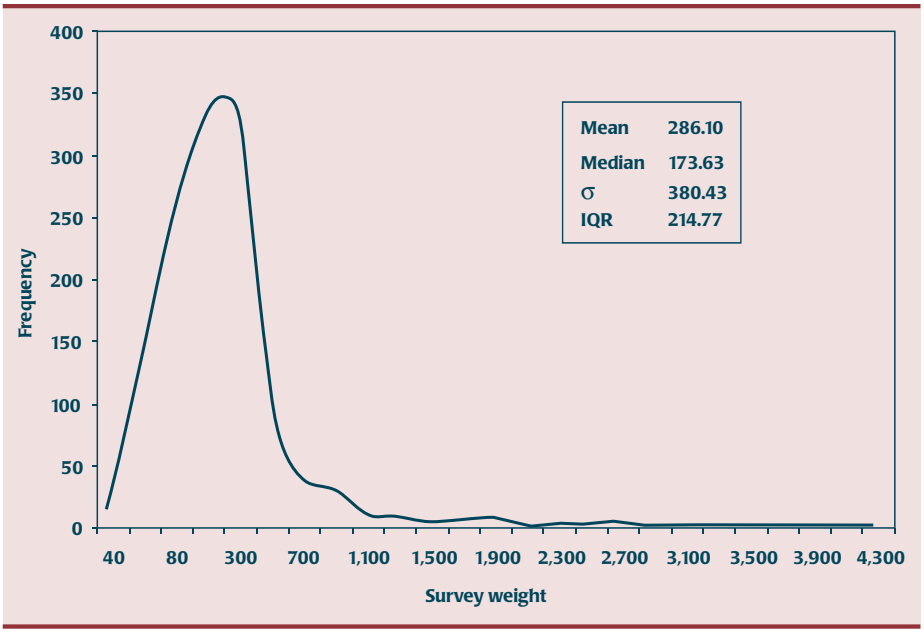
were primarily due to the inclusion of the survey weights not unusual values of *x* and *y*. Both observations have survey weights well over 1,000. More importantly, these observations are now having undue influence on the parameter estimates as evidenced by the large *dffit* values (see Figure 4). Large survey weights are not necessarily problematic in and of themselves. The problem here occurs because the distribution of the survey weights in the NLSCY is skewed to the right. Figure 5 shows the smoothed distribution of the longitudinal survey weights for males.

The weights range from a low of 20 to a high of just over 4,000. There are very few observations with extreme weights; when paired with a less than perfect fit, these observations can have a large influence on the parameter estimates. The survey weights are determined by the sampling design and as such are geographically determined. Fifteen observations in the weighted model with a *dffits* value greater than 1.0 or *Cook's D* greater than 0.3 were defined as influential and were excluded. All of the influential cases were from large urban centres in large provinces (Ontario, Quebec, or BC) and had survey weights in excess of 1,000. It was not possible to get individual profiles for the dropped cases due to confidentiality reasons. However, it was possible to identify variables associated with large survey weights. Province (Ontario, Quebec and BC) and large urban centres were clearly associated with weights in the top 5th percentile. In boys, cases with survey weights in the top 5th percentile were 32 more times likely to come from a large urban centre as compared to a small urban centre and 42 times more likely to come from a large urban centre as compared to a rural area. In addition, cases with large survey weights were four times more likely to have been born within Canada/US, and four times as likely to come from a single parent household. The survey weight was not associated with the outcome of vocabulary

The impact of removing the influential observations was far greater in the weighted model than in the unweighted model (see Table 3). The maximum studentized residual dropped from 7.5 to 4.2 in the weighted model whereas the maximum studentized residual changed from 3.8 to 4.2 in the unweighted model, supporting the contention that the influential observations were driven primarily by the sampling weight as opposed to unusual values of *x* and *y*.

The impact of removing the influential observations relative to both the size and variability of the estimate can also be seen by comparing the parameter estimates in the full model to the parameter estimates from the final model, excluding the influential observations. Removing the influential observations altered the parameter estimates for geographic area (rural), immigration status, size for gestational age, single parent status, parental age, education, income, social support, and family dysfunction by greater than one standard deviation (see Table 4). The influential observations had resulted in an overestimation of the effects of geographic area, single parent status,

FIGURE 5
Smoothed distribution of the longitudinal survey weight for boys age 2 to 3 years from cycle 1 of the NLSCY



parental age, income, and family dysfunction and an underestimation of the effects of size for gestational age, education and social support.

Discussion

The basic weight for the NLSCY was derived from the LFS. The sampling strategy for the LFS was designed to equally represent various geographical areas. The use of the LFS to identify households with children resulted in non-contiguous strata and sometimes only one cluster within a stratum. The resulting distribution of sampling weights is markedly skewed with a few cases with extremely large weights relative to the rest of the sample. A skewed weight distribution will result in a weighted analysis being more inefficient than an unweighted analysis.[8] However, as demonstrated in this paper, another important, and often overlooked, implication of a skewed weight distribution is the potential of heavily weighted observations to have an undue influence on the parameter estimates. Although methods such as robust estimation[9] and down weighting[8,10] have been suggested for the treatment of influential clusters in survey data, all of these approaches require access to stratum and cluster indicator variables and assume that there is more than one cluster per stratum. This is not always the case in surveys such as the National Longitudinal Survey of Children and Youth (NLSCY) and the National Population Health Survey (NPHS) and is the reason that Statistics Canada recommends the use of bootstrap procedures to correctly estimate the variance.

Skewed weight distributions are a characteristic of not only the NLSCY but also the NPHS. Both are extremely rich data sources that can be used to help answer a number of population health questions with important implications for guiding policy development over the next 10 years. A comparison of the weighted and unweighted analyses is a useful tool during preliminary model building. Large unexpected differences between the parameter estimates from the weighted and unweighted models can help flag important missing covariates, model misspecifiation, or undetected interaction. All analyses should include a detailed influence analysis. Observations that are having an undue influence on the weighted model fit should be excluded from the final model. Given the size of the original sample, removal of a few observations should not have a meaningful impact on the generalizability of the results. Failure to do so could result in erroneous conclusions, and in the case of the NLSCY, measures of association unduly influenced by the experience of children from large cities in large provinces. The standard error of the parameter estimates in the final model should be calculated using bootstrap variance estimation with the replicate weights and code available from Statistics Canada.

References

1. Korn, EL, Graubard, BI. *Analysis of health surveys*. New York (NY): John Wiley & Sons Inc.; 1999.
2. Human Resources Development Canada, Statistics Canada. *National Longitudinal Survey of Children and Youth: User's handbook*. Ottawa: Minister of Industry; 1999.
3. Fox, J. *Applied regression analysis, linear models, and related methods*. Thousand Oaks, (CA): Sage Publications, Inc.; 1997.
4. Kleinbaum, DG, Kupper, LL, Muller, KE, Nizam, A. *Applied regression analysis and multivariable methods*. 3rd ed. Pacific Grove (CA): Brooks/Cole; 1998.
5. Statistics Canada. *The Labour Force Survey*. Ottawa (ON): Statistics Canada; 1996.
6. Dunn, LM. *Peabody picture vocabulary test – revised*. American Guidance Service; 1981.
7. Belsley, DA, Kuh, E, Welsch, RE. *Regression diagnostics. Identifying influential data and sources of collinearity*. New York (NY): John Wiley & Sons; 1980.
8. Zaslavsky, AM, Schenker, N, Belin, TR. Down-weighting influential clusters in surveys: Application to the 1990 post enumeration survey. *J Am Stat Assoc* 2001;96:858-69.
9. Gwet, JP, Rivest, LP. Outlier resistant alternatives to the ratio estimator. *J Am Stat Assoc* 1992; 87:1174-82.
10. Hidiroglou, MA, Srinath, KP. Some estimators of a population total from simple random samples containing large units. *J Am Stat Assoc* 1981;76: 690-95.

The estimation of heritability for twin data based on concordances of sex and disease

Hongzhuan Tan, Mark Walker, France Gagnon, and Shi Wu Wen

Abstract

Heritability is an important measure in chronic disease epidemiology. Almost all developed methods of heritability estimation for dichotomous outcomes in twin data are based on concordance of monozygous (MZ) and dizygous (DZ) twins. However, most existing twin registries, which provide a unique and efficient opportunity to assess the putative genetic basis of diseases, do not have zygosity information. We developed a method that can be used to estimate the heritability for twin data with no information on zygosity. The only conditions on using this method are that the studied disease incidence is not strongly related to sex, and the distribution of zygosity is in accordance with Weinberg's rule. Using asthma twin data which has histological confirmed zygosity, we compared the results of our method with Holzinger's formula. The heritability of asthma was 24.88% (95%CI 21.98% - 27.78%) and 29.83% (95%CI 22.28% - 37.38%) estimated by our method and Holzinger's formula respectively. We conclude that our new method can be used in the estimation of heritability with large twin register data with no zygosity information available.

Key words: *heritability, sex, twin registers, zygosity, Weinberg's rule*

Introduction

Heritability (h^2) could be defined as the ratio of additive genetic variance (S_g^2) to phenotypic variance ($S_p^2 = S_g^2 + S_e^2$ [environmental variance]).[1] Heritability estimates provide baseline information about the genetic architecture of a chronic disease or dichotomous trait. The heritability calculated for a disease can be used in genetic consultation, as a reference in clinical diagnosis, and can help in decision-making processes such as setting priorities for research funding allocation or prevention program development.[2–5] Genetic variance and herit- ability can be estimated in several ways. The most direct method is to obtain an estimate of S_e^2 by making a number of homozygous lines from the population, crossing them in pairs to reconstitute individual heterozygotes, and measuring the phenotypic variance within each heterozygous genotype, and thus estimating S_e^2. The S_g^2 can then be obtained by S_g^2 subtracted from the value of S_p^2 in the original population.

Other estimates of genetic variance can be obtained by considering the genetic similarities between relatives from family data plus population data, twin data plus population data, or only from twin data.[1] The commonly used methods of heritability estimation for dichotomous variables in twins include Falconer's method and Holzinger's method. Falconer's method is based on the difference between concordance rate of studied disease among monozygous (MZ) and dizygous (DZ) twin pairs and incidence rate in general population.[6] Holzinger's formula,[7–8] is based only on the observed concordance rate of studied disease among monozygous (MZ) and dizygous (DZ) twin pairs. The commonly used method of heritability estimation for continuous variables in twins is Nichol's method, the correlation of variables in each class of twins instead of the concordance rate ($h^2 = 2 (r_{MZ} - r_{DZ}) / r_{MZ}$).[9] All of these methods require information of zygosity for every twin pair.

There are many large twin registries in the world that can be rich resources for chronic disease genetic epidemiologic studies.[10–12] Such registries usually provide unbiased data, since they are representative of an entire population. However, histologically determined zygosity information is not available in most of these large computerized data registries. As a result, Holzinger's formula cannot be used. Sofaer and Holloway[13] have reported a heritability estimation method that is based on the concordance of twin pairs, the proportion of same sex (SS) pairs among all concordance affected pairs, and the prevalence of the studied disease in the general population. Sofaer & Holloway's method estimates the heritability (h^2) by an iterative procedure using a range of possible values for h^2, calculating the corresponding P_E (the expected proportion of concordant affected pairs like-sexed) at each iteration and choosing as h^2 that value of h^2 which make the P_E closest to P_O (the observed proportion of concordant affected pairs like-sexed). In addition to its complex calculation, Sofaer & Holloway's method needs the general population data about the studied disease prevalence and uses data on concordance twin pairs only, resulting in reduced efficiency. Moreover, Sofaer & Holloway's method is restrictive to

Author References

Hongzhuan Tan, Shi Wu Wen, OMNI Research Group, Department of Obstetrics & Gynecology, University of Ottawa, Faculty of Medicine, Ottawa, Ontario, Canada, *and* Ottawa Health Research Institute Clinical Epidemiology Program, *and* School of Public Health, Central South University, Changsha, Hunan, P. R. China
Walker Mark, OMNI Research Group, Department of Obstetrics & Gynecology, University of Ottawa, Faculty of Medicine, Ottawa, Ontario, Canada,
France Gagnon, Department of Epidemiology and Community Medicine, University of Ottawa, Faculty of Medicine, Ottawa, Ontario, Canada
Correspondence: Shi Wu Wen, OMNI Research group, Department of Obstetrics & Gynecology, The Ottawa Hospital, 501 Smyth Rd, Box 241, Ottawa, ON Canada, K1H 8L6. Fax: (613) 739-6266; E-mail: swwen@ohri.ca

diseases with prevalence ranging from 0.1% to 10%. In this paper, we introduce a simpler and more effective method that estimates heritability using twin data without confirmation of zygosity.

Modification of Holzinger's method

From Holzinger's method of heritability estimation we deduced the formula making it suitable for twin data without zygosity information. Holzinger's formula of heritability estimate from twin data is as follows:

$$h^2 = (C_{MZ} - C_{DZ}) / (1 - C_{DZ}) \quad (1)$$

here, the h^2 is heritability, C_{MZ} is proband concordance rate of disease in MZ twin pairs, and C_{DZ} is proband concordance rate of disease in DZ twin pairs. The proband concordance rate is given as $C_p = (C + C') / (C + D + C')$, where C is the total number of + + pairs, C' is the number of + + pairs in which the two affected members have been ascertained independently, and D is the number of + - pairs. When ascertainment is complete, $C = C'$, so that $C_p = 2C / (2C + D)$[8]. On the basis of Holzinger's formula, we set:

C_{OS}: proband concordance rate of disease in opposite sex (OS) twin pairs

C_{SS}: proband concordance rate of disease in same sex (SS) twin pairs

N_{SS}: number of SS twin pairs

N_{OS}: number of OS twin pairs

Pr (DZ | SS): probability of being DZ twin pairs in SS twin pairs

Pr (MZ | SS): probability of being MZ twin pairs in SS twin pairs

Given that all OS twin pairs are DZ, suppose that the occurrence of the studied dichotomous trait is not related to sex, the proband concordance rate of the disease in OS twin pairs is equal to the proband concordance rate of the disease in DZ twin pairs ($C_{DZ} = C_{OS}$). According to Weinberg's rule,[8] the probability of DZ in SS twin pairs equals to the ratio of N_{OS} and N_{SS}, namely:

$$\Pr(DZ \mid SS) = N_{OS}/N_{SS}$$

Thus

$$\Pr(MZ \mid SS) = 1 - \Pr(DZ \mid SS) = 1 - N_{OS}/N_{SS}$$

$$C_{SS} = \Pr(DZ \mid SS) * C_{DZ} + \Pr(MZ \mid SS) * C_{MZ}$$

And

$$C_{MZ} = (C_{ss} - \Pr(DZ \mid SS) * C_{DZ}) / \Pr(MZ \mid SS) = (C_{SS} - N_{OS}/N_{SS} * C_{OS}) / (1 - N_{OS}/N_{SS})$$

So we have:

$$h^2 = (C_{MZ} - C_{DZ}) / (1 - C_{DZ}) = ([(C_{SS} - N_{OS}/N_{SS} * C_{OS}) / (1 - N_{OS}/N_{SS})] - C_{OS}) / (1 - C_{OS}) \quad (2)$$

Using formula (2), we can calculate the heritability only with the proband concordance rate of disease in SS and OS twin pairs, and the number of SS and OS twin pairs.

If the sample size is large, we can calculate the standard error of heritability S (h^2) and the confidence interval (CI of h^2 with the formula[14]:

$$(3)^*$$

$$95\%\,CI \text{ of } h^2 = h^2 \pm 1.96S\,(h^2) \quad (4)$$

Here, N_{MZ} and N_{DZ} are the number of MZ and DZ twin pairs respectively. When zygosity information is not available, we can have $C_{MZ} = (C_{SS} - N_{OS}/N_{SS} * C_{OS}) / (1 - N_{OS}/N_{SS})$, $C_{DZ} = C_{OS}$, $N_{DZ} = 2N_{OS}$, $N_{MZ} = N_{SS} - N_{OS}$ by applying Weinberg rule.

In this formula, we have assumed that the occurrence of the studied dichotomous trait is not strongly related to sex and the SS dizygotic twin pairs are approximately equal in number to OS twin pairs according to the Weinberg rule.[8]

If we know the proportion of DZ ($\Pr_{(DZ)}$) in all twin pairs in a population, we can use this $\Pr_{(DZ)}$ to calculate the heritability, combined with the $\Pr_{(OS)}$ (the observed proportion of OS twin pairs in all twin pairs), $\Pr_{(SS)}$ (the observed proportion of SS twin pairs in all twin pairs), and the C_{all} (concordance rate of the studied dichotomous trait in all twins). Under this situation, the second condition is not required, and we can modify the formulae for C_{MZ} and h^2. We have:

$\Pr_{(MZ)}$ (the proportion of MZ twin pairs in all twin pairs) = $1 - \Pr_{(DZ)}$

Because:

$$\Pr_{(DZ)} * C_{DZ} + \Pr_{(MZ)} * C_{MZ} = C_{all}$$

So:

$$C_{MZ} = (C_{all} - \Pr_{(DZ)} * C_{DZ}) / \Pr_{(MZ)} = (C_{all} - \Pr_{(DZ)} * C_{OS}) / (1 - \Pr_{(DZ)}) \quad (5)$$

$$h^2 = (C_{MZ} - C_{DZ}) / (1 - C_{DZ}) = ([(C_{all} - \Pr_{(DZ)} * C_{OS}) / (1 - \Pr_{(DZ)})] - C_{OS}) / (1 - C_{OS}) \quad (6)$$

In order to estimate the heritability with formula 2, we need to know C_{OS}, C_{SS}, N_{OS}, and N_{SS}. If we use formula 6, we need to know $\Pr_{(DZ)}$, C_{OS}, and C_{all}. All these variables are usually available in large register databases.

Application to asthma twin data set

We used published data[15] of twins with asthma with histologically determined zygosity as an example to validate our method. The study sample included 3,807 twin pairs. Of these, 1,232 pairs are monozygous female (MZF), 566 pairs are monozygous male (MZM), 751 pairs are dizygous female (DZF), 352 pairs are dizygous male (DZM), and 906 pairs are dizygous opposite sex (DZO). There were three types of twin pairs according to the distribution of asthma: both twins had asthma (D-D), one of the twins had asthma (D-N), and neither twin had asthma (N-N). The results are listed in Table 1. Based on data in Table 1, we calculated the heritability (h^2) of asthma using Holzinger method and the method presented in this paper, and compared the results. Only the D-D and D-N twin pairs were used in the calculation of concordance rate.

Using Holzinger method:

Concordance rate in monozygous twins (C_{MZ}):

$C_{MZ} = (106*2) / ([106*2] + 249) = 0.4599$

Concordance rate in dizygous twins (C_{DZ}):

$C_{DZ} = (63*2) / ([63*2] + 421) = 0.2303$

* $S(h^2) = \sqrt{\left[\frac{1 - C_{MZ}}{(1 - C_{DZ})^2}\right]^2 \cdot \frac{C_{DZ}(1 - C_{DZ})}{N_{DZ}} + \left(\frac{1}{1 - C_{DZ}}\right)^2 \frac{C_{MZ}(1 - C_{MZ})}{N_{MZ}}}$

Heritability $h^2 = (C_{MZ} - C_{DZ}) / (1 - C_{DZ}) = (0.4599-0.2303) / (1-0.2303) = 29.83\%$

The standard error of heritability S (h^2) = 0.0385

95% confidence interval (CI) of h^2: 22.28% - 37.38%

Using our newly developed method

$N_{ss} = 2901$
$N_{os} = 906$

Concordance rate in same sex twins (C_{ss}):
$C_{ss} = (137*2) / ([137*2] + 467) = 0.3698$

Concordance rate in opposite sex twins (C_{os}):
$C_{os} = (32*2) / ([32*2] + 203) = 0.2397 = C_{DZ}$

Concordance rate in monozygous twins (C_{MZ}):
$C_{MZ} = (C_{SS} - N_{OS}/N_{SS} * C_{OS}) / (1 - N_{OS}/N_{SS}) = (0.3698 (906/2901)*0.2397) / (1-906/2901) = 0.4289$

Heritability $h^2 = (C_{MZ} - C_{DZ}) / (1 - C_{DZ}) = (0.4289 - 0.2397) / (1- 0.2397) = 24.88\%$

The standard error of heritability
S (h^2) = 0.0148

95% confidence interval (CI) of h^2:
21.98% - 27.78%

Discussion

Large unbiased twin registries provide a unique opportunity to verify the role of genetics in chronic disease etiology, which is a crucial step before extensive and expensive genetic epidemiologic research is undertaken. Unfortunately, because information on zygosity is often not available in these data, available methods cannot be used to calculate heritability from the twin registry data. On the other hand, information on infant sex is almost always available in registry data. Heritability estimates based on sex and disease concordance, such as the method introduced in this paper, will facilitate chronic disease genetic epidemiologic studies, particularly in the use of existing large twin registries.

In the formula of heritability estimation, none of the parameters are related to disease prevalence. The concordance (C) is related to the genetic effect only, and not related to the prevalence. As shown in Table 1, the heritability estimation is just related to the D-D and D-N, and is not related to the N-N. The N_{OS}/N_{SS} is a ratio: it is determined by Weinberg rule and the ratio of N_{MZ} to N_{DZ}, and is not related to the prevalence of the disease. If the prevalence changes, it only influences the number of N-N, and the number of N-N is not related to our formula. Therefore, our method has no restriction to the prevalence of disease. Our method requires that SS dizygotic twin pairs are approximately equal in number to OS twin pairs according to Weinberg rule.[8] Although there are some variations in the ratio of SS to OS in DZ twin pairs, the ratio is very close to the ratio indicated by Weinberg rule.[16,7] Our method is therefore applicable to many chronic diseases or dichotomous traits, although it is limited to diseases or dichotomous traits with no strong sex effect. If the studied disease or dichotomous trait is strongly related to sex, C_{DZ} will not equal to C_{OS}, and the formula will be incorrect. In addition, since the method is based on affection status (i.e., disease concordance), the estimates will be more reliable for diseases with a clear distinction between affected and unaffected individuals.

Our method is simpler to calculate than Sofaer & Holloway method,[13] and, because it uses both concordance twin pairs and discordant twin pairs, it is more efficient. Sofaer & Holloway method is limited to diseases of the prevalence range 0.1% to 10%, whereas our method theoretically has no restriction in disease prevalence.

As illustrated in the example, using disease concordance in same sex and opposite sex twin pairs, our method yielded an estimation of heritability for asthma of 24.88% (95% CI 21.98% 7.78%), which is quite similar to the heritability estimated by Holzinger method with histologically confirmed zygosity data for the twins (29.83%, 95% CI 22.28% 7.38%). Although the studied disease (asthma) is somewhat related to sex, the C_{OS} (0.24) estimated by our method is very close to the C_{DZ} (0.23) estimated by Holzinger method, and the S(h^2) from our method (0.0148) is lower than the S(h^2) from Holzinger method (0.0385), indicating that our method is comparable in terms of precision. Furthermore, the two assumptions in our method, namely the disease occurrence not being strongly related to sex and, as many DZ twin pairs in SS twin pairs as the number of OS twin pairs, can be satisfied in many chronic diseases or dichotomous traits. As a result, our method can find a broad application.

TABLE 1
Number of twin pairs with asthma by zygosity and sex

	Sex and Zygosity Group								
	MZF	MZM	MZ	DZF	DZM	DZO	DZ	SS	OS
D-D	67	39	106	19	12	32	63	13	32
D-N	185	64	249	136	82	203	421	467	203
Sub-total	252	103	355	155	94	235	484	604	235
N-N	980	463	1,443	596	258	671	1,525	2,297	671
Total	1,232	566	1,798	751	352	906	2,009	2,901	906
Concordance rate	0.42	0.55	0.46	0.22	0.23	0.24	0.23	0.37	0.24

(Hopper JL, Genet Epidemiol, 1990, 7 (4):277 898)

Acknowledgments:

Dr. Tan is a University of Ottawa International Fellow. Dr. Wen is a Canadian Institute for Health Research New Investigator, and Dr. Walker is an Ontario Ministry of Health Career Scientist.

References

1. Falconer DS, Mackay TFC. *Introduction to quantitative genetics.* Fourth Edition, Harlow, Longman Publishers, 1996.
2. Meng W, Lu H, Cai R, Jiang F, Jiang Q, Lin W, et al. A study on the genetic epidemiology of hepatocellular carcinoma. *Zhonghua Liu Xing Bing Xue Za Zhi* 2002;23(6):438–40.
3. Swan L, Birnie DH, Padmanabhan S, Inglis G, Connell JM, Hillis WS. The genetic determination of left ventricular mass in healthy adults. *Eur Heart J* 2003;24(6):577–82.
4. Rietveld MJ, Hudziak JJ, Bartels M, Van Beijsterveldt CE, Boomsma DI. Heritability of attention problems in children: Cross-sectional results from a study of twins, age 3-12 years. *Am J Med Genet* 2003;117B(1): 102–13.
5. North KE, Howard BV, Welty TK, Best LG, Lee ET, Yeh JL, et al. Genetic and environmental contributions to cardiovascular disease risk in American Indians: the strong heart family study. *Am J Epidemiol* 2003;157(4):303–14.
6. Falconer DS. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet*. 1965;29:51–76.
7. Levitan M. *Textbook of human genetics*, Third Edition, New York, Oxford University Press, 1988.
8. Emery AEH, Rimoin DL. *Principles and practice of medical genetics*. Churchill Livingstone Inc., New York, 1983.
9. Nichol RC. The National Merit Twin Study, in *Methods and goals in human behaviour genetics.* Academic Press, Vandenburg, Chicago, 1965.
10. Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. *Nature reviews/genetics* 2002;3:872–82.
11. National Center for Health Statistics. *1995 997 Matched multiple birth data set.* NCHS CD-ROM Series 21, No. 12. Hyattsville, Maryland, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2000.
12. Busjahn A. Twin registers across the globe: What's out there in 2002? *Twin Res* 2002;5(5):v–vi
13. Sofaer JA, Holloway SM. Heritability estimation from concordance twin pairs alone. *Acta Genet Med Genellol* 1986;35:49–60.
14. Vogel-Motulsky. *Human genetics problems and approaches*. Springer-Verlag Berlin Heidelberg, New York, 1982.
15. Hopper JL, Hannah MC, Macaskill GT, Mathews JD. Twin concordance for a binary dichotomous trait: Ø. A bivariate analysis of hay fever and asthma. *Genet Epidemiol* 1990;7(4):277–89.
16. Vlietinck R, Derom C, Derom R, Van den Berghe H, Thiery M. The validity of Weinberg's rule in the East Flanders Prospective Twin Survey (EFPTS). *Acta Genet Med Gemellol* 1988;37(2):137-41.
17. Husby H, Holm NV, Gernow A, Thomsen SG, Kock K, Gurtler H. Zygosity, placental membranes and Weinberg rule in a Danish consecutive twin series. *Acta Genet Med Gemellol* 1991;40(2):147–52.

Breast cancer trends in Manitoba: 40 years of follow-up

Alain A Demers, Donna Turner, Daojun Mo, and Erich V Kliewer

Abstract

This study reports a comprehensive array of breast cancer statistics for Manitoba for a 40-year period. Data from the Manitoba Cancer Registry were combined with the provincial population-based registration file to determine trends in breast cancer incidence, prevalence and mortality rates, as well as survival and the probability of being diagnosed with breast cancer in the next 10 years. The age-standardized incidence rate of breast cancer increased by 0.99/100,000 women per year over the 40 years of follow-up (69.6/100,000 women in 1960, 109.9/100,000 women in 1999). Mortality rates peaked in 1986 (35.7/100,000 women), while the 1999 mortality rate (26.0/100,000 women) was almost comparable to the 1960 rate (22.4/100,000 women). No significant trend in mortality rate was observed over the 40-year period. The 5-year prevalence rate of breast cancer increased by 8.6/100,000 women per year. Between 1960-64 and 1995-99, 5-year survival increased from 0.62 to 0.86. The probability of being diagnosed with breast cancer in the next 10 years increased the most for women 60 years of age. The breast cancer burden in Manitoba is rapidly evolving mainly because of the increasing incidence and the better survival of cases.

Keywords: *breast neoplasms; incidence; prevalence; registries; vital statistics*

Introduction

Breast cancer is the most frequently diagnosed cancer in women around the world. Its incidence is increasing in most industrialized countries; even in some areas of China, where the risk of breast cancer is low, the incidence rates have risen by more than 50% between 1972 and 1994.[1-5] Worldwide, new cases of female breast cancer accounted for 18% of all cancer cases in 1980, 19% in 1985, 21% in 1990, and 22% in 2000.[3-6] Breast cancer is also a leading cause of death attributed to cancer. It was estimated that 174,000 women died from breast cancer in developed countries in 1990, which represents 17% of all deaths due to cancer in women.[7]

Although incidence and mortality rates tend to be readily available for many industrialized countries, historical information on other statistics such as prevalence and survival are rarely accessible. These data are key to cancer surveillance and the monitoring of cancer control activities. With the increasing incidence of breast cancer and the decreasing mortality from the disease, health care planners require accurate information on prevalent cases in order to plan the allocation of services for a rapidly evolving situation. A recent estimate by Pisani et al.[8] indicated that the 5-year prevalence rate of breast cancer in high-income countries was 392/100,000 women in 1990. However, information on prevalence remains scarce for most jurisdictions and where available, it is estimated mainly from incidence and mortality rates.[9]

Canada, like many other countries, routinely publishes information on breast cancer incidence and mortality, but less on prevalence and survival. On a national level, Health Canada has estimated that 21,200 women (30% of all female cancer cases) will be diagnosed with breast cancer in 2004, which makes it the most frequent in women.[10] Among the Canadian provinces, Manitoba had the highest average age-standardized incidence rate of breast cancer between 1990 and 2000 (105.8/100,000 women) followed by Nova Scotia (103.6/100,000 women). Between 1989 and 1999, Manitoba has had an average age-standardized mortality rate (28.3/100,000 women), which was comparable to the overall Canadian rate (28.9/100,000 women).[11]

While these data illustrate the significance of breast cancer as a public health concern in Manitoba, key questions remain about the historical trends in incidence and mortality as well as the burden of the disease on the population and the health care system as measured by prevalence and survival. In particular, are these relatively high rates a new phenomenon? What kind of prognosis (survival) can Manitobans expect, and has it improved? How do these trends translate into prevalence? The current study addresses these questions by examining trends of breast cancer incidence, mortality and 5-year prevalence rates as well as relative survival and the probabilities of being diagnosed with breast cancer from 1960 to 1999.

Method

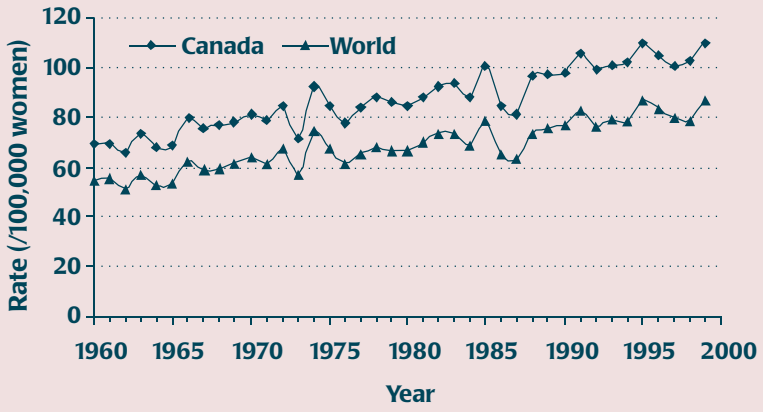
Data sources and population

The Manitoba Cancer Registry, which is housed at CancerCare Manitoba, was established in 1937 and became population-based in 1956. Multiple sources of ascertainment of incident cases are used, including physician

Author References

Alain A Demers, Donna Turner, Erich V Kliewer, CancerCare Manitoba, Department of Epidemiology and Cancer Registry, Winnipeg, Manitoba *and* Department of Community Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada
Daojun Mo, CancerCare Manitoba, Department of Epidemiology and Cancer Registry, Winnipeg, Manitoba
Correspondence: Alain A Demers, CancerCare Manitoba, Epidemiology and Cancer Registry, 675 McDermot Avenue, Winnipeg, Manitoba, R3E 0V9;
Fax: (204) 786-0628, E-mail: alain.demers@cancercare.mb.ca

FIGURE 1
Age-standardized incidence rates of breast cancer (/100,000 women) in Manitoba between 1960 and 1999 (rates were standardized to the 1991 Canadian and World population)



notifications, pathology and hematology reports, and hospitalization, mortality, and autopsy records. For every case, the Cancer Registry includes information on diagnosis according to the ICD-9 code (ICD-10 since 2002), age at diagnosis and date of diagnosis, tumour grade, tumour morphology and vital status. The Manitoba Vital Statistics department provides information on mortality. In examining cases registered from 1991 to 1995, the North American Association of Central Cancer Registries estimated the Manitoba Cancer Registry to be 95% complete in ascertaining all cancer cases.[12]

Incident cases of malignant breast cancer (ICD-9 174) diagnosed among Manitoba residents and recorded in the registry from January 1, 1960 to December 3l, 1999 were included in the analysis. For women who had multiple primary breast cancers, only information on the first event reported was considered. Breast cancer cases could have been diagnosed with other primaries than breast cancer before their first diagnosis of breast cancer. Information used for the calculation of cause-specific mortality rates came from the Vital Statistics Agency of the Government of Manitoba. Prevalence and relative survival were determined using death information from the Manitoba Cancer Registry in combination with the Manitoba Population Registration file maintained by Manitoba Health. The population file includes coverage start and end dates (death, emigration) for all Manitobans registered for the provincial Health Insurance Plan. Two-year and 5-year prevalence were defined as the number of women who had been diagnosed with breast cancer in Manitoba and who were still alive and resident of the province two and five years after diagnosis. Total prevalence was based on those women who were alive and still residing in Manitoba at the end of the study period. Information on emigration was obtained from the Manitoba Health Population Registration file. Follow-up information (deaths, migration) was available for 13,665 (73%) of the 18,596 women diagnosed with breast cancer over the study period. Of the women with coverage information, 7,388 (54%) were still living in Manitoba on January 1, 2000, 5,767 (42%) died within the study period, and 510 (4%) emigrated.

Analyses

Incidence and mortality rates were standardized to the age distribution of both the 1991 Canadian population[10] and the world

TABLE 1
Age-specific incidence rates of breast cancer (/100,000 women) in Manitoba between 1960 and 1999 by period of diagnosis

	Age groups					
Period of diagnosis	40–49	50–59	60–69	70–79	≥80	Standardized rate[b]
1960–64	119	158	163	230	248	69
1965–69	133	162	198	259	293	77
1970–74	136	184	245	236	283	82
1975–79	144	172	242	259	336	84
1980–84	140	211	266	306	311	91
1985–89	126	201	276	373	345	93
1990–94	122	236	329	391	378	102
1995–99	138	261	335	366	355	105
Annual percent change in incidence rates[a]	1960–99: 0.06	1960–99: 1.45	1960–99: 1.80	1960–81: 0.66 1981–88: 5.91 1988–99: –0.99	1960–99: 0.89	1960–99: 1.14

[a] Annual percent changes (APC) were determined by Joinpoint analyses; APCs were calculated by age group using single calendar year, although rates are presented by 5-year periods; periods for which Joinpoint indicated that there was a significant change in the trend are shown.

[b] Rates were standardized to the 1991 Canadian population.

TABLE 2
Age-specific and age standardized mortality rates of breast cancer (/100,000 women) in Manitoba between 1960 and 1999 by period

Period	Age groups					Standardized rate[b]
	40–49	50–59	60–69	70–79	≥80	
1960–64	29	62	77	87	149	26
1965–69	40	77	81	98	158	77
1970–74	29	69	93	129	179	30
1975–79	28	65	95	103	151	28
1980–84	27	64	91	119	174	29
1985–89	29	66	93	135	183	31
1990–94	24	53	97	135	179	28
1995–99	28	59	86	125	205	29
Annual percent change in mortality rates[a]	1960–99: -0.73	1960–99: -0.55	1960–99: 0.26	1960–99: 0.95	1960–99: 0.66	1960–99: 0.08

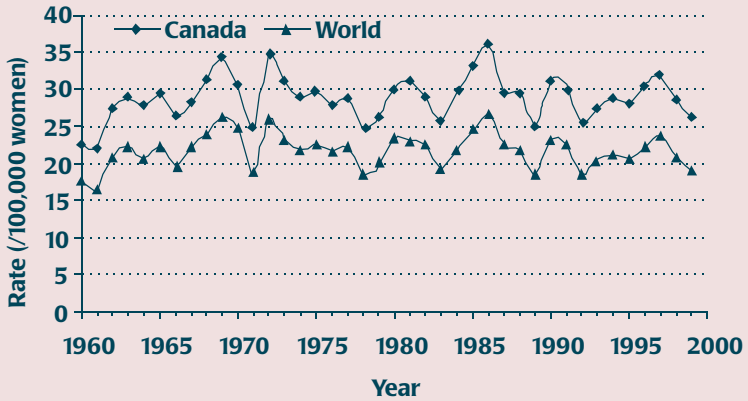
[a] Annual percent change (APC) were determined by Joinpoint analyses; APC were calculated by age group using single calendar year, although rates are presented by 5-year periods; periods for which Joinpoint indicated that there was a significant change in the trend are shown.

[b] Rates were standardized to the 1991 Canadian population.

population.[13] They were also standardized to the European[13] and 2000 US[14] populations for discussion purposes. Crude prevalent rates, instead of adjusted rates, were presented as suggested by Zanetti et al.[9] These authors pointed out that crude prevalence is a more useful and informative tool in determining the utilization of health services. Trends and annual percent change (APC) were determined using Joinpoint Regression Program (v2.7) distributed by the National Cancer Institute (NCI) Surveillance, Epidemiology, and End Results (SEER) division.[15] The APC was calculated using the log-linear model, where the APC is equal to $100 * (e^m - 1)$ and m is the estimated slope of the regression line.

Population and death counts used in the calculation of relative survival and probabilities of being diagnosed with breast cancer were obtained from Statistics Canada.[16] A SAS program developed by Paul Dickman that uses the Ederer II method[18] for calculating the expected number of deaths was used to calculate the relative survival probabilities.[17] The endpoint for survival was death, emigration (when available) or the end of the study period (December 31, 1999). Cancers diagnosed by death certificates only (0.15%) or from death reports from unofficial sources (0.01%) were excluded from the relative survival analyses. The probabilities of being diagnosed with breast cancer were calculated using DEVCAN 5.2.[19,20] Data management was performed using SAS (v8.2).

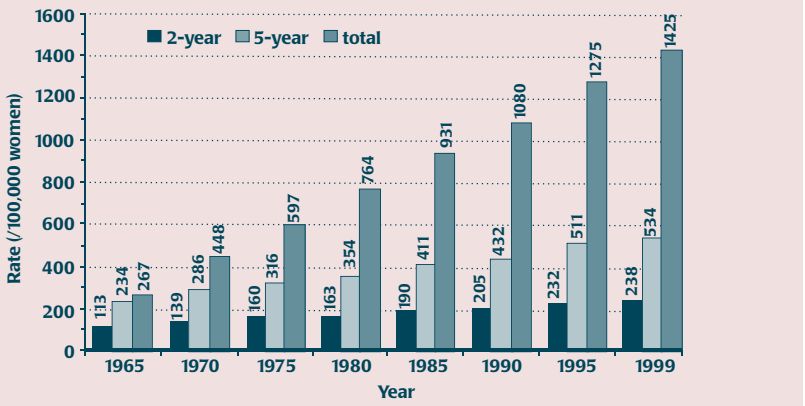
FIGURE 2
Age-standardized mortality rates from breast cancer (/100,000 women) in Manitoba between 1960 and 1999 (rates were standardized to the 1991 Canadian and the World population)



Results

A total of 18,596 women were diagnosed at least once with breast cancer in Manitoba between January 1, 1960 and December 31, 1999. Over this time period, 6.1% of women diagnosed with breast cancer were under 40 years of age, 17.2% were aged 40–49 years, 21.3% were aged 50–59 years, 23.5% were aged 60–69 years, 20.5% were aged 70–79 years and 11.5% were 80 years of age and older. The information of diagnosis age was missing for 0.06% of cases.

FIGURE 3
Two-year, 5-year and total prevalence rates of breast cancer in Manitoba between 1965 and 1999 by year



Although there were fluctuations in the annual incidence and mortality rates as a result of the relatively small Manitoba population, the age-standardized incidence rate increased by 0.99/100,000 women per year ($p < 0.0001$) over the 40-year follow-up (Figure 1). Overall, the incidence rate was 58% higher in 1999 than in 1960. The changes were not the same across all age groups (Table 1). An increase in incidence was observed in women older than 50 years of age ($p < 0.0001$), but mainly in the 60–69 year age group. The Joinpoint analysis indicated that the incidence rate declined in the 70–79 age group between 1988 and 1999, although it is too early to consider these results as a definitive trend.

The age-standardized mortality rate from breast cancer remained stable between 1960 and 1999 (0.038/100,000 women per year, $p = 0.35$) (Figure 2). Age-specific mortality rates fluctuated in the different age groups, most likely as a consequence of low frequencies, although a decrease was noted in those under 60 years of age ($p < 0.05$) and an increase occurred in those 70 years of age and over ($p < 0.03$) (Table 2). The number of deaths per year ranged from 84 to 220, with the highest number of deaths occurring in 1997, although the highest annual rate was observed in 1986.

The 2-year, 5-year and total prevalence rates all increased over time (Figure 3). Between 1965 and 1999, the increase per 100,000 women per year was 3.4 for two-year prevalence, 8.6 for 5-year prevalence and 33.4 for total prevalence. The age-specific 5-year prevalence rate increased the most among women 60–69 and 70–79 years of age although it rose in all women over 50 years of age (Table 3).

The five-year relative survival rate from breast cancer improved from 0.62 in 1960–64 to 0.86 in 1995–99 periods (Figure 4). Over the same time periods, the probability of being diagnosed with breast cancer for the first time in the next 10 years was relatively constant for women 30 and 40 years of age (Table 4). Women aged 60 years experienced the highest probability increase; it went from 1.6% in 1960–64 to 3.2% in 1995–99.

TABLE 3
Age-specific and total five-year prevalence rates of breast cancer (/100,000 women) in Manitoba between 1960 and 1999 by five-year group

Year	Age groups					Crude rate
	40–49	50–59	60–69	70–79	≥80	
1965	400	654	679	909	1,105	234
1970	548	731	803	1,095	1,214	286
1975	542	822	1,043	984	1,151	316
1980	599	754	1,039	1,190	1,403	354
1985	562	903	1,207	1,459	1,472	411
1990	470	853	1,211	1,733	1,540	432
1995	502	1,036	1,496	1,810	1,797	511
1999	515	1,147	1,545	1,691	1,737	534
Annual percent change in prevalence rates[a]	1964–80: 2.17 1980–90: -2.57 1990–99: 1.66	1964–99: 1.40	1964–99: 2.12	1964–79: 0.65 1979–91: 4.30 1991–99: -0.84	1964–99: 1.58	1964–99: 2.35

[a] APC: Annual percent change (APC) were determined by Joinpoint analyses; APC were calculated by age group using single calendar year, although rates are presented by 5-year periods; periods for which Joinpoint indicated that there was a significant change in the trend are shown.

TABLE 4
Probability (%) of Manitoba women without breast cancer being diagnosed with breast cancer for the first time in the next decade by period

	Period							
Age	1960–64	1965–69	1970–74	1975–79	1980–84	1985–89	1990–94	1995–99
30	0.4	0.4	0.4	0.5	0.4	0.4	0.4	0.5
40	1.1	1.2	1.3	1.3	1.3	1.2	1.3	1.4
50	1.5	1.6	1.8	1.8	2.0	1.9	2.3	2.5
60	1.6	1.9	2.2	2.3	2.5	2.7	3.1	3.2
70	1.9	2.2	2.1	2.3	2.7	3.3	3.6	3.4

Discussion

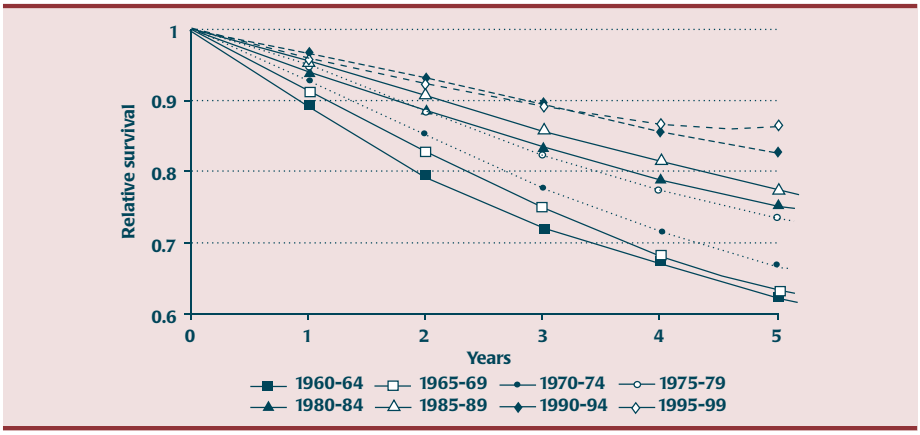
The age-standardized incidence rate of breast cancer increased by 0.99/100,000 women per year in Manitoba between 1960 and 1999, which represents an overall relative increase of 58%. This observation is consistent with results reported in other western countries.[2] In 1990, the Manitoba incidence rate was one of the highest (77/100,000 women, standardized to the World population) among the countries assessed by Parkin et al.[3] Manitoba's 1990 incidence rate (106/100,000 women, standardized to the European population) was comparable to Belgium's rate and would have ranked the third highest among 15 European countries evaluated by Black et al.[21]

Within the North American context, the 1995–99 Manitoba incidence rate (114.3/100,000, standardized to the 2000 US population) was lower than the rate reported by SEER for the same time period (136.7/100,000 women of all races).[22] In the US, the incidence rate leveled off in 1987 for white women and seems to be decreasing in women 75 years of age and older.[23,24] A similar trend may be happening since 1995 in Manitoba for 70–79-year-old women, although it is too soon to make a decisive conclusion. With the exception of 1986, Manitoba's rates have been consistently higher than the Canadian average since 1984.[11] Manitoba had an average of five more breast cancer cases per year (/100,000 women) than Canada between 1984 and 2000.

Over the 40 years of follow-up, no significant change was observed in Manitoba's mortality rate. In other words, no inflection point was found using Joinpoint. However, there were differences by age with younger women showing a decrease in mortality while older women experienced an increase in mortality. Reports of national trend data indicate that there has been a decrease in mortality since the late 1980s.[10] Manitoba's mortality rates have been comparable to the Canadian average since at least 1984.[11] Compared to the rest of the world, Manitoba's 1990 mortality rate (23/100,000 women, standardized to the world population) would have ranked 17th out of 63 countries evaluated by Levi et al.[25] When standardized to the European population, the 1990 mortality rate (33/100,000 women) would have ranked Manitoba fifth highest among the 15 European countries evaluated by Black et al.[21] The 1995–99 breast cancer death rate was 28.8/100,000 women (all races) in the US,[22] which is lower than the Manitoba 1995–99 death rate (32.3/100,000 women, standardized to the US 2000 population).

The 2-year, 5-year and total breast cancer prevalence rates have all shown increases since the 1960s. People who survive five years or more are often considered to be "cured" and, as such, may be the prevalence measure of most interest.[8] The 5-year breast cancer prevalence rate increased by 2.3% per year in Manitoba between 1964 and 1999. The increase was mainly apparent in women 60 years of age and older. The aging population, the increase in Manitoba residents, and the longer survival of women diagnosed with breast cancer are the principal factors responsible for this trend.[26] In Manitoba, the proportion women 60 years of age and older increased from 12.4% of the female population in 1960 to 19.4% in 1999. In the same time period, the number of women residing in the province increased from 417,500 to 577,669 individuals. These changes, together with an increasing incidence rate, had the effect of increasing number of incident breast cancer cases during a time when the relative survival was also increasing and thus leading to a marked increase in prevalence. In 1990, Manitoba's 5-year prevalence rate (438/100,000 women) was higher than the average 5-year prevalence rate reported in the world's high-income countries (392/100,000 women).[8] However, when the comparison group is restricted to European countries, Manitoba's 1990 5-year prevalence rate is comparable to these countries.[9] European

FIGURE 4
Cumulative relative survival of women diagnosed with breast cancer in Manitoba by period of diagnosis



rates varied between 275/100,000 women in Greece to 543/ 100,000 women in Sweden. The average was 411/100,000 women for the European Union, slightly higher than the Mani gold10toba rate.

Relative survival from breast cancer increased between 1960–64 and 1995–99. Two jumps are apparent, one between 1965–69 and 1975–79 and one between 1985–89 and 1990–94. These increases may be due to improvement in treatment and diagnosis practices, but also, to some extent, to a lead-time bias introduced by population-based screening. The study by Welch et al[27] using data form the National Cancer Institute supports this phenomenon. The authors reported that an improvement in five-year survival over time bears little relationship to changes in cancer mortality, but appears to be related to earlier detection of tumours. However, it should be kept in mind that causes of death listed on death certificates are not always accurate. The Manitoba breast-screening program was implemented in 1995 and elective breast screening has been available since the 1970s. The percentage of Manitoba women 50 to 69 years old having a mammogram increased from 43% in 1990–94 to 68% in 1995–99 and thus, some of the recent increase in relative survival may be the result of changes in screening. Similar trends in survival improvement have been reported in Europe,[28] the United States,[29] and Asia.[30] The five-year relative survival of women diagnosed in Manitoba in 1985–89 (77%) was comparable to observations made in Europe, where the five-year relative survival ranged from 58% (Slovakia) to 81% (Sweden) over the same time frame.[28] The five-year relative survival was 86.2% (all races) in the US in 1992–98[22] and 83% in Canada in 1992–94,[31] which are in agreement with the Manitoba five-year relative survival (86.4%) for the 1995–99 time period.

The probability of being diagnosed with breast cancer in the next decade increased over the 40 years of follow-up for women 50, 60 and 70 years of age. The 1995–1999 probabilities reported are consistent with contemporaneous observations made by Morris et al.[32] in California and Health Canada.[10] Morris et al. reported a risk of developing breast cancers in the next 10 years of 3.4% for women 60 years of age living in California in 1993–97, while Health Canada reported a risk of 3.1% for women 60–69 years of age living in Canada in 2000. In the present study, the risk of developing breast cancer in the next 10 years for a woman aged 60 years was 3.2% in 1995–99.

In summary, the incidence and five-year prevalence rates of breast cancer have significantly increased in Manitoba between 1960 and 1999. Over the 40-year period the mortality rate from breast cancer has been fairly constant, although there are indications that there has been a decrease since the late 1980s. The five-year cumulative relative survival has improved significantly over the four decades. These statistics suggest that the burden of cancer in Manitoba is high but is in the range of observations made in Canada nationally and in other westernized countries.

Acknowledgments

Donna Turner is supported in part by a Senior Research Fellowship from the Canadian Institutes for Health Research (CIHR).

References

1. Jin F, Devesa SS, Chow W-H, et al. Cancer incidence trends in urban Shanghai, 1972-1994: An update. *Int J Cancer* 1999; 83(4):435–40.
2. Parkin DM. Breast cancer in Europe: epidemiology and forecasts. *Eur J Oncol* 1999; 2:45–64.
3. Parkin DM, Pisani P, Ferlay J. Estimates of the worldwide incidence of 25 major cancers in 1990. *Int J Cancer* 1999;80(6):827–41.
4. Parkin DM, Läärä E, Muir CS. Estimates of the worldwide frequency of sixteen major cancers in 1980. *Int J Cancer* 1988; 41:184–97.
5. Parkin DM, Pisani P, Ferlay J. Estimates of the worldwide incidence of eighteen major cancers in 1985. *Int J Cancer* 1993; 54:594–606.
6. Parkin DM, Bray FI, Devesa SS. Cancer burden in the year 2000. The global picture. *Eur J Cancer* 2001;37(Suppl 8):S4–66.
7. Pisani P, Parkin DM, Bray F, Ferlay J. Estimates of the worldwide mortality from 25 cancers in 1990. *Int J Cancer* 1999;83(1): 18–29.
8. Pisani P, Bray F, Parkin DM. Estimates of the world-wide prevalence of cancer for 25 sites in the adult population. *Int J Cancer* 2002; 97(1):72–81.
9. Zanetti R, Micheli A, Rosso S, Sant M. The prevalence of cancer: a review of available data. *Tumori* 1999;85:408–13.
10. National Cancer Institute of Canada. *Canadian cancer statistics 2004.* Toronto, Canada: National Cancer Institute of Canada; 2004 April.
11. Health Canada. Disease surveillance on-line. http://www.hc-sc.gc.ca/pphb-dgspsp/dsol-smed/ index.html 2004 Access date: May 2004.
12. Chen VW, Wu XC, Andrews PA. *Cancer in North America, 1991-1995*, volume one: incidence. Sacramento, CA: North American Association of Central Cancer Registries; 1999 April.
13. IARC. In: *Cancer incidence in five cotinents volume III*. Lyon, France: WHO, International Agency for Research on Cancer (IARC); 1976. p. 456.
14. U.S. Cancer Statistics Working Group. *United States cancer statistics: 2000 incidence.* Atlanta (GA): Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2003.
15. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med 2000*;19:335–51.
16. Statistics Canada. *Canadian socio-economic information management system* (CANSIM). 2004. http://datacentre.chass.utoronto.ca Access date: July 2002.
17. Dickman P. *Population-based cancer survival analysis 2002*. http://www.pauldickman.com/teaching/index.html Access date: June 2002.
18. Ederer F, Heise H. *Instruction to IBM 650 programmers in processing survival computations. Methodological note No. 10*. Bethesda MD: End Results Evaluation Section, National Cancer Institute; 1959.
19. Fay MP, Pfeiffer R, Cronin KA, Le C, Feuer EJ. Age-conditional probabilities of developing cancer. *Stat Med* 2003; 22(11):1837–48.

20. National Cancer Institute – Statistical Research and Applications Branch. DevCan: Probability of Developing or Dying of Cancer Software, Version 5.2. 2004. http://srab.cancer. gov/devcan Access date: June 2004.

21. Black RJ, Bray F, Ferlay J, Parkin DM. Cancer incidence and mortality in the European Union: Cancer registry data and estimates of national incidence for 1990 (published erratum appears in *Eur J Cancer* 1997 Dec; 33(14):2440). *Eur J Cancer* 1997; 33(7): 1075–1107.

22. Ries L, Eisner M, Kosary C, et al. *SEER Cancer Statistics Review, 1973-1999.* http://seer.cancer.gov/csr/1973_1999/ 2002 Access date: June 2004.

23. Howe HL, Wingo PA, Thun MJ, et al. Annual report to the nation on the status of cancer (1973 through 1998), featuring cancers with recent increasing trends. *J Natl Cancer Inst* 2001;93(11):824–41.

24. Ries LA, Wingo PA, Miller DS, et al. The annual report to the nation on the status of cancer, 1973-1997, with a special section on colorectal cancer. *Cancer* 2000; 88:2398–2424.

25. Levi F, Lucchini F, Negri E, La Vecchia C. Worldwide patterns of cancer mortality, 1990-1994. *Eur J Cancer Prev* 1999; 8(5):381–400.

26. Rothman KJ, Greenland S. Measures of disease frequency. In: Rothman KJ, Greenland S, editors. *Modern Epidemiology*. Philadelphia: Lippincott Williams and Wilkins; 1998. p. 29–46.

27. Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? *JAMA* 2000; 283(22):2975–78.

28. Quinn MJ, Martinez-Garcia C, Berrino F. Variations in survival from breast cancer in Europe by age and country, 1978-1989. EUROCARE Working Group. *Eur J Cancer* 1998;34(14 Spec No):2204–11.

29. Frost F, Tollestrup K, Hunt WC, Gilliland F, Key CR, Urbina CE. Breast cancer survival among New Mexico Hispanic, American Indian, and non-Hispanic white women (1973-1992). *Cancer Epidemiol Biomarkers Prev* 1996;5(11):861–66.

30. Chia KS, Du WB, Sankaranarayanan R, Sankila R, Seow A, Lee HP. Population-based cancer survival in Singapore, 1968 to 1992: an overview. *Int J Cancer* 2001; 93(1):142–47.

31. Statistics Canada. Cancer survival statistics. 2004. http://www.statcan.ca/english/freepub/84-601-XIE/survival.htm Access date: June 2004.

32. Morris CR, Wright WE, Schlag RD. The risk of developing breast cancer within the next 5, 10, or 20 years of a woman's life. *Am J Prev Med* 2001;20(3):214–18.

Smoker preference for "elastic cigarettes*" in the Canadian cigarette market

Michael O Chaiton, Neil E Collishaw, and Aaron J Callard

Abstract

Elastic cigarettes are characterized by yields of constituents that increase proportionally faster than smoke volume as cigarettes are smoked more intensely. Elasticity may function to overcome physical limitations in increasing puff volume during nicotine-seeking behaviour. The purpose of this study was to determine if there are elastic cigarettes in the Canadian cigarette market, and to determine smoker preference for elastic cigarettes. Elasticity was calculated for 115 brands in the Canadian filtered cigarette market for puff volumes of 44 and 56 ml. Puff volumes, nicotine and tar deliveries were obtained from earlier published documents.[1] Sales data were used as a proxy for smoker preference. Ordinary least squares regression was used to determine the association of sales and elasticity in the Canadian cigarette market. The cigarette brands ranged from a mean elasticity value of 1.21 to 0.67. Of the 115 Canadian cigarette brands tested, 23 brands had a mean elasticity value significantly over 1.00, making them elastic. After adjusting for brand, the average elastic cigarette sold an average of 361 million cigarettes while an inelastic cigarette sold 89.5 million cigarettes ($p < 0.0001$). The difference in sales between elastic and inelastic cigarettes was independent of tar yield and filter type. Elasticity was not associated with tar yield ($p = 0.2734$). There are elastic cigarettes in the Canadian cigarette market and the results suggest a possible smoker preference for elastic cigarettes. Utilizing elasticity may be valuable in the development of future harm reduction strategies.[2]

Key words: *compensation; elasticity; nicotine; tobacco*

Introduction

Elastic cigarettes are characterized by yields of constituents that increase proportionally faster than smoke volume as cigarettes are smoked more intensely. Elasticity describes a non-linear response-to-effort relationship between smoker effort, as measured by puff volume, and cigarette delivery, and was explored as a method for smokers to get more nicotine out of low delivery cigarettes by facilitating compensatory smoking.[2,3,4] The concept of elasticity can also be applied to tar, total particulate matter, or tar/nicotine ratio.[3,4]

Elasticity is achieved as a summary function of a number of design features including ventilation, pressure drop, packing density, chemistry of tobacco, physical properties of tobacco, air permeability of the cigarette papr, and the composition of the cigarette paper.[3] The two most important are percent ventilation and rod porosity. These design features can be adjusted by manufacturers to obtain different smoking properties, including the property of elasticity.

British American Tobacco (BAT) scientists found that smokers had received a fairly constant amount of nicotine irrespective of the International Organization of Standardization (ISO) rating of nicotine delivery, either compensating down with higher nicotine cigarettes or compensating up when the level was low.[5] They concluded that nicotine was both the signal (as the sensation of impact) and the driving force for the compensating behaviour, noting methods of compensation such as covering the ventilation holes of the cigarette in light and mild cigarettes.[2]

Studies of human smoking behaviour found that smokers were limited in their ability to compensate by the amount of effort that was needed to draw on the cigarette.[6,7,8] Elastic cigarettes were proposed within the tobacco industry as a way of reducing the "reward for effort" relationship.[3] A small increase in puff volume could bring a proportionately larger reward of nicotine, producing a more satisfying cigarette. Elastic cigarettes may make it easier for smokers to control the dose of nicotine, either increasing or decreasing the desired dose.

Gray and Kozlowski suggest that cigarettes were produced to be more elastic as tar and nicotine levels were reduced.[2] They suggest that the elastic cigarette facilitated nicotine compensation, was easier to learn to smoke, and was potentially more addicting than non-elastic cigarettes.[2,8] Elastic cigarette are, conceivably, a more potent nicotine delivery device than un-elastic cigarettes. However, the effects of elastic cigarette have not been well examined outside of the tobacco industry.

The purpose of this study is to assess the elasticity of cigarettes in the Canadian market and the association of elasticity and sales in Canada. It would be expected that, in general, cigarette brands would be inelastic, as tobacco industry studies have found

Author References

Michael O Chaiton, University of Toronto and Physicians for a Smoke-Free Canada, Ottawa, Canada
Neil E Collishaw, Aaron J Callard, Physicians for a Smoke-Free Canada, Ottawa, Canada
Correspondence: Michael O Chaiton, University of Toronto, 33 Russell Ave. Toronto, Ontario, Canada M5S 2S1; Fax: (416) 595-6068;
E-mail: michael.chaiton@utoronto.ca

* The term "elastic cigarettes" is synonymous with "compensatible cigarettes".

that the relative yield of nicotine is un- affected by changes in puff volume.[9] On the other hand, if elastic cigarettes were better nicotine delivery devices, then sales of elastic cigarettes should be higher than those of inelastic cigarettes, reflecting a preference for a more adjustable nicotine system.

Method

Elasticity formula

The Brown and Williamson formula for "normalized elasticity" is used for comparison between brands of differing deliveries.[4] The formula first defines a hypothetical cigarette that has a linear, or "rigid," response to increasing puff volume. The predicted linear relationship between puff volume and per-puff delivery can be imagined as a regression line that passes through the origin. Elasticity is a numerical expression of the relationship between observed and predicted delivery in response to an increase in puff volume. To determine elasticity, the observed delivery per puff over a certain realistic range, such as 44 ml to 56 ml, is divided by the predicted delivery per puff (D2/P2 in the formula below). The predicted delivery per puff is based on the assumption that delivery per puff varies in linear fashion with puff volume.

The formula for elasticity, as described by Gonterman[4] can be summarized as follows:

$$\text{Elasticity} = [(D_2/P_2)/(D_1/P_1)] / (V_2/V_1)$$

Where:

- V_1 = puff volume 1 (the calibrating puff volume, e.g., 44 ml)
- V_2 = puff volume 2 (the experimental puff volume, e.g., 56 ml)
- D_1 = delivery per cigarette in mg at puff volume 1
- D_2 = delivery per cigarette in mg at puff volume 2
- P_1 = number of puffs per cigarette at puff volume 1
- P_2 = number of puffs per cigarette at puff volume 2

An elastic cigarette will deliver proportionately more nicotine than would be expected for a given increase in puff volume. In numerical terms, a number above 1.00 indicates that the increase of nicotine (or other chemical of interest) is proportionally greater than the change in the puff volume, or "elastic".[10] More crudely stated, an elastic cigarette will deliver "more bang for your suck."

Elasticity calculation

Puff volumes, puff number, and nicotine deliveries for 115 brands manufactured by Imperial Tobacco Canada (ITL), Rothmans, Benson & Hedges (RBH) and JTI Macdonald (JTI) were obtained from the 1996 report "Determination of Cigarette Yields under Realistic Conditions" by WS Rickert of Labstat Inc.[1] The 1996 report was initially performed for Health Canada to determine realistic smoking deliveries. Puff volumes of 44 ml (V_1) and 56 ml (V_2) were chosen in conjunction with the Labstat report and tobacco company documents as being realistic levels of normal and intense smoking.[1,4,6] Other than puff volume, all other variables were held constant (2 seconds puff duration, 26 seconds puff interval, unobstructed ventilation). The Labstat report described the mean value and the standard deviation of 10 repeated measures for nicotine and puff number for each brand tested. The Brown and Williamson normalized elasticity formula was then used to calculate elasticity.

A Monte Carlo simulation was used to determine the confidence intervals for elasticity.[3] Values for the parameters of nicotine and puff number for each brand were randomly drawn from their respective probability distributions and the elasticity was calculated. This calculation was repeated 10,000 times. The 2.5th and 97.5th percentiles were taken from the distribution of the repetitions of calculations of the elasticity to represent the endpoints of a 95% confidence interval.

A number of assumptions were made for this approach. First, the mean distribution of the nicotine yield and puff number data in the Labstat report were both assumed to be approximately normal in order to draw from a normal probability function. Repeated measurement of the yield and puff number should form a normal curve around the true value and so this distribution was used. Second, the variation in puff volume levels in the Labstat report was assumed to be insignificant. Because no standard deviation was given for this value, it had to be assumed that puff volume was a constant and, if any variation in volume did occur, it would be too small to be of consequence.

FIGURE 1
Histogram of elasticity values of Canadian cigarettes in 1996. Elasticity calculated from Rickert, 1996

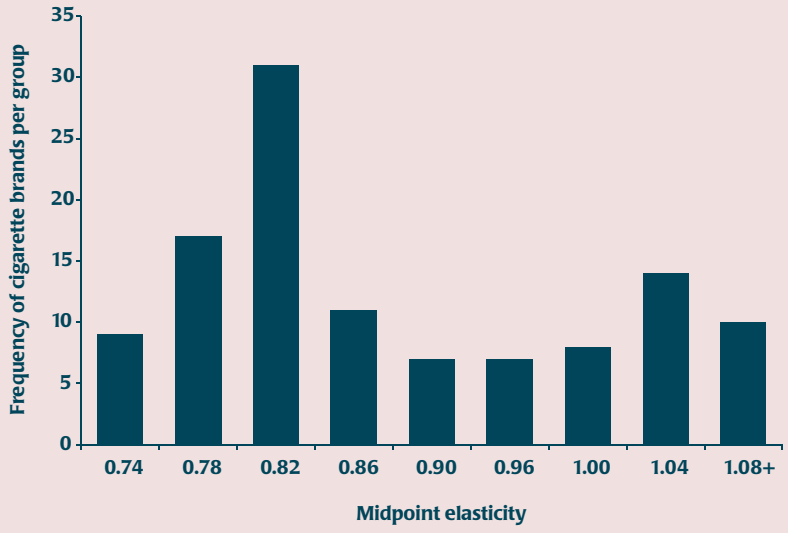


TABLE 1
Average ISO nicotine, tar values and sales of elastic, inelastic, and linear response cigarettes in the Canadian cigarette market, 1996

Elasticity	N	ISO Nicotine Mean (95% Confidence interval)		ISO Tar Mean (95% Confidence Interval)		Sales (millions of cigarettes sold) Mean (95% Confidence Interval)		Median sales
Inelastic	83	1.0	(0.9,1.0)	10.3	(9.3,11.2)	140.4	(100.4,180.4)	95.1
Linear	9	0.9	(0.5,1.3)	9.4	(4.9,13.9)	772.7	(161.6,1383.7)	290.7
Elastic	23	1.1	(1.0,1.2)	11.7	(10.5,12.9)	1202.4	(590.9,1813.8)	719.3

Note: Elasticity calculated from Rickert, 1996. Tar and nicotine data from Rickert, 1996. Sales data from Health Canada.[11] Inelastic (less than 1.00 elasticity); Linear (not significantly different than 1.00 elasticity); Elastic (greater than 1.00 elasticity.

Third, the reported standard deviation of the nicotine yield and puff number was assumed to be not significantly different from the true deviation, and so the reported standard deviation was used in the Monte Carlo simulation.

As the elasticity value itself is not as important as determining whether elasticity was significantly above 1.00 (elastic) or significantly below 1.00 (inelastic), the z-test was used to determine whether the value of elasticity was significantly different from 1.00 at an alpha = 0.05 level. Cigarette brands were classified as either not significantly different than 1.00 (linear), significantly greater than 1.00 (elastic) or significantly less than 1.00 (inelastic). All subsequent uses of elasticity refer to the categorized variable.

Association with sales

Ordinary least squares (OLS) regression was used to determine the association of sales and elasticity in the Canadian cigarette market. Two models were created. The first controlled for differences in manufacturer and differences in tar yield and the second examined the effects of brand. Sales data was taken from the report of Canadian tobacco manufacturers of 1996 sales data to Health Canada.[11] The standard yield of tar was as determined by methods described by the ISO, included in Rickert (1996).[1] The logarithm of sales was used in the regression to normalize the distribution of sales.

To control for brand, a process of matching was used. Many cigarette brands come in sets, differing only in the type of filters used. It was assumed that identical brands would have similar properties related to branding, advertising, distribution, and tar and nicotine yield. For instance, "Players Extra Light" King Size and "Players Extra Light" Regular were matched and coded as being the same brand. Cigarette brands that did not have a pair were excluded, leaving 72 types of cigarettes (36 pairs) for the analysis. The association of elasticity and filter type was tested using the chi-square test to ensure the independence assumed in the model.

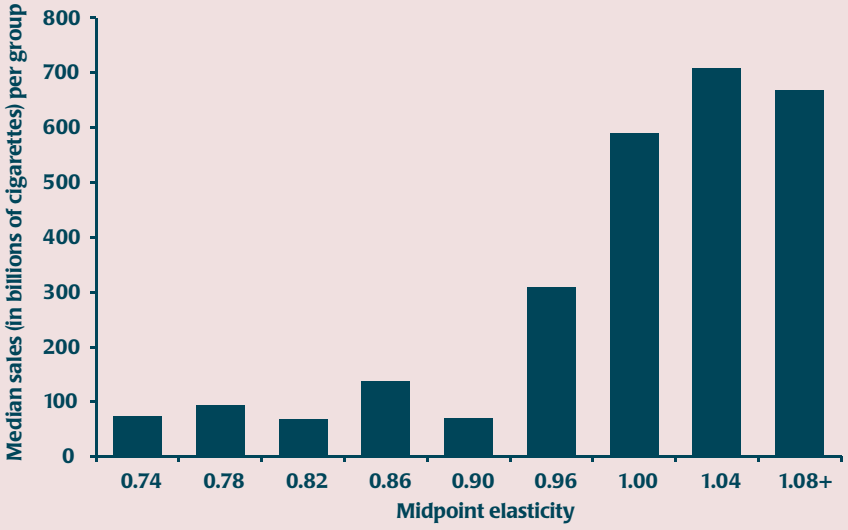
Association with tar yield

OLS regression was used to determine the association of tar yield and elasticity. The standard yield of tar was as determined by methods described by the ISO included in Rickert (1996).[1]

Results

The cigarette brands ranged from a mean elasticity value of 1.26 to 0.70 (Figure 1) with a median value of 0.84. Of the 115 Canadian cigarette brands tested, 23 brands (20%) had a mean elasticity value significantly over 1.00, making them elastic. Eighty-three brands were inelastic (72%) and there were 9 brands (8%) with a linear increase of nicotine proportional to the puff volume. However, inelastic cigarettes accounted for only 25% (11.6 billion cigarettes) of all cigarettes sold in the sample compared to 15% (7.0 billion cigarettes) for linear delivery cigarettes and 60% (27.7 billion cigarettes) for elastic cigarettes. Elastic cigarettes had marginally higher average tar and nicotine values, and much higher average sales per brand than inelastic cigarette brands (Table 1). Sales per brand are fairly constant until the elasticity approaches 1.00 (Figure 2). While Japan Tobacco International and Rothmans Benson and Hedges had elastic brands (4/25, 6/50), most elastic cigarettes were produced by Imperial Tobacco Limited (13/40), accounting for 57% of all elastic cigarettes (chi-square 5.83, p value = 0.014).

FIGURE 2
Median sales by elasticity group in the Canadian cigarette market, 1996. Sales data via Health Canada.[11] Elasticity calculated from Rickert, 1996

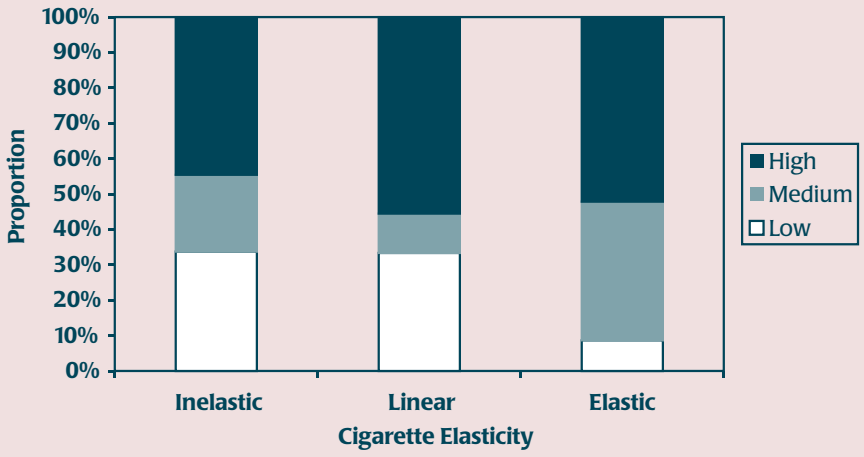


After controlling for manufacturer and tar yield, elasticity was significantly associated with cigarette sales (F value 31.18; p value $< .001$). Holding manufacturer and tar yield constant, the difference between inelastic and elastic cigarettes was associated with a seven-fold increase in cigarette sales (Table 2). There were significant differences in sales between cigarettes with a linear response and inelastic cigarettes (t value = -4.43; $p < 0.001$) and between elastic and inelastic cigarettes (t value = -6.93; $p < 0.001$); however, there was no significant difference between linear and elastic cigarettes (t value = -0.33; $p = 0.74$).

Controlling for brand, the relationship between elasticity and sales was still strong (F value = 8.47; p value = < 0.001). Elastic cigarettes were associated with a four-fold increase in sales compared to inelastic cigarettes of the same brand (Table 2). Filter type was not associated with elasticity (chi-square test = 0.99; p value = 0.91) or sales (F value = 0.12; p value = 0.88).

There was no overall significant association between elasticity and ISO tar value (F value = 1.31, p value = 0.2734). Elastic cigarettes were distributed in all levels of tar yield and were not particularly associated with lower tar yield. In fact, there were few "low tar"cigarettes (ISO tar values ≤ 8) among elastic cigarettes brands compared to the relative proportion among inelastic or linear response cigarette brands (Figure 3).

FIGURE 3
Proportion of high ($\leq$ 13 mg tar), medium ($>$ 8 mg and $<$ 13 mg tar), and low ($\leq$ 8 mg tar) ISO tar values in inelastic, linear response and elastic cigarette brands on the Canadian cigarette market, 1996



Discussion

There were elastic cigarettes in the Canadian cigarette market. It was not surprising that the market would include some brands that were elastic as there were a range of cigarette brands with a variety of cigarette design and manufacturing techniques available on the market; nevertheless, a majority, 72%, of cigarette brands on the Canadian cigarette market were inelastic. However, elastic cigarette brands accounted for 60% of cigarettes sold, reflecting greater sales of elastic brands compared to inelastic brands. After controlling for the variety of effects associated with brand, the association between sales and elasticity continued to be strong.

It would be expected that any smoker preference for elastic cigarettes would be manifested particularly in lower tar delivery cigarettes where the access to extra nicotine would be desired by many smokers.[3] However, there was no association between ISO tar yields and elasticity. Elastic cigarette brands were found in the full range of ISO delivery levels from full flavour to light and mild. In higher delivery brands, elastic cigarettes may allow the smoker to subconsciously adjust their nicotine intake to allow for more accurate titration of nicotine, either up or down, to the level that they are seeking. One explanation is that it is more important for the cigarette to be flexible in delivery than to deliver a larger amount of nicotine. Elasticity may allow a varying nicotine requirement to be more easily and less consciously self-regulated. BAT scientists, at one time, contemplated a variable delivery cigarette that "adjusts to varying need (satisfaction) of the consumer during the course of a day" and which would have operated by changes in puff volume.[12]

The uniformity in price, tobacco type and flavourings in the Canadian tobacco market may allow for design differences in cigarettes to be reflected in sales that may be harder to demonstrate in other markets.[13] The results of this study are similar to a Philip Morris study, which found that a regression model including advertising, extractable nicotine, and other cigarette design features statisti-

TABLE 2
Sales of elastic, inelastic and linear response cigarettes adjusted for the effect of manufacturer and ISO tar yield or brand in the Canadian cigarette market, 1996. Adjusted least square mean sales (millions of cigarettes) (95% Confidence Interval)

Elasticity	Controlling for Manufacturer and ISO tar yield		Controlling for brand	
Inelastic	69.2	(53.2, 90.0)	89.5	(69.8, 114.7)
Linear	432.0	(196.9, 947.9)	340.0	(170.6, 677.8)
Elastic	503.6	(304.4, 833.2) (n=115)	360.7	(228.2, 570.2) (n=72)

Note: Elasticity calculated from Rickert, 1996. Sales data from Health Canada.[11] Inelastic (less than 1.00 elasticity); Linear (not significantly different than 1.00 elasticity); Elastic (greater than 1.00 elasticity.

cally explained over 95% of the variability in market share between brands.[14,15] The cigarette design features in the study, including extractable nicotine, were described as balancing the smoke properties in American-style blended cigarettes produced with Turkish, Oriental, and Burley tobacco. It is conceivable that elasticity might perform an equivalent role to extractable nicotine in Canadian-style Virginia tobacco cigarettes or alternatively, elasticity may act as a proxy for extractable nicotine content.

This study is limited by having only two puff volumes. A larger range or a different set of puff volumes may provide a different interpretation for both the overall elasticity of the cigarette and any reaction of the smoker. Elasticity values may also have changed since the data was collected, as well as sales figures, and are unlikely to reflect the current cigarette brands. A longitudinal study of the relationship of sales and elasticity would be preferable. Further research is needed to clarify the role of elasticity in smoker preference in the Canadian cigarette market. To do this, more detailed information on cigarette design is needed; specifically, the variables of percent ventilation and rod porosity, which are the two main factors in elasticity, are needed for public examination. While this study has pointed to a clear relationship between elasticity and cigarette sales, it is based on a single cross-sectional measurement. From current data it is not possible to determine neither the causality of the relationship nor the direction of causation.

If properly studied and understood, elasticity could be utilized for effective regulation of cigarette design that would limit the ability of smokers to titrate nicotine. In conceivable harm reduction strategies, preventing elastic cigarettes could allow for regulation for true low nicotine yield cigarettes or for cigarettes that would be less able to maintain addictions. It may be possible to regulate cigarette design cigarettes that are less toxic or less appealing to the user. Banning ventilation of cigarettes might be an option to consider in this regard as a step in reducing the elasticity of cigarettes and providing a more accurate baseline of cigarette deliveries.[2,8]

On the other hand, it is also conceivable that, in fact, there is a public health benefit to elastic cigarettes by facilitating nicotine dosing, while minimizing toxin delivery. However, since elasticity requires an increase in puff volume, reducing toxin delivery would be more sensibly achieved by having a cigarette that could deliver nicotine at low puff volumes, without need to increase volume and consequently increase delivery of other constituents. A further possibility is that smokers of elastic cigarettes may require fewer cigarettes per day than smokers of inelastic cigarettes; however, the sales difference between inelastic and elastic cigarette suggest that this may not be true on a population basis. More detailed research is needed to understand the influence of elasticity on the life history of the cigarette smoking to determine if association between sales and elasticity is due to differences in uptake, use, cessation, or another factor.

Conclusion

There are elastic cigarettes in the Canadian cigarette market and these cigarettes enjoy a prominent sales advantage. Understanding the reasons behind this advantage is important to future attempts at regulating cigarette design. Elasticity may allow consumers to more easily adjust their smoking behaviour to seek additional nicotine.

Acknowledgements

Funding generously provided by Canadian Tobacco Control Research Initiative.

References

1. Rickert WS *Determination of cigarette yields under realistic conditions*. Prepared for Health Canada by Labstat Inc. 1996.

2. Gray N and Kozlowski L. More on the regulation of tobacco smoke: how we got here and where next. *Annals of Oncology*, 2003; 14:353–57.

3. Schneider-W B&W. *Elasticity of cigarettes*. 10 Sep 1992. Bates: 575251611–575251643. http://tobaccodocuments.org/bw/956817.html

4. Gonterman R. *Elasticity data/399*. 07 Apr 1992. Bates: 583221740–583221742. http://tobaccodocuments.org/bw/964632.html

5. Pangritz D. *Smoke elasticity*. "Document 04000752". http://tobaccodocuments.org/health_canada/04000752.html

6. Dunn PJ, Freiesleben ER. *The use of the freiri slave smoker to investigate changes in smoking behaviour: Part I. 03 Mar 1975. Bates: 650007446–650007479. http://tobaccodocuents.org/bw/17418.html*

7. Dunn PJ, Freiesleben ER. *The use of the Freiri Slave Smoker to investigate changes in smoking behaviour: Part II*. 25 Mar 1975. Bates: 650007421–650007445. http://tobaccodocuments.org/bw/17417.html

8. Kozlowski LT, O'Connor RJ. Cigarette filter ventilation is a defective design because of misleading taste, bigger puffs, and blocked vents. *Tobacco Control* 2002;11 (Suppl 1): i40–i50

9. Thornton R, Bat. *The effect of changes in puff volume on smoke chemistry. Report No. Rd. 384–R. 660328*. Rd&E Southampton. 28 Mar 1966. Bates: 570512244–570512269B. http://tobaccodocuments.org/bw/955868. html

10. *Imperial Tobacco Ltd. Research & Development Division Montreal. Progress Report: July 93 – December 93.* available at www.tobaccopapers.org at pages 402415168–217.

11. Fact sheets on market share of canadian tobacco manufacturers. 1997 available at www.smoke-free.ca.

12. *Implications and activities arising from correlation of smoke ph with nicotine impact, other smoke qualities, and cigarette sales*. 08 May 1996. Bates: 2041197302– 2041197334. http://tobaccodocuments.org/pm/2041197302-7334.html

13. Tong HS. *The pharmacology of smoke – dose nicotine a review of current literature*. 07 Oct 1976. Bates: 80192103–80192113. http://tobaccodocuments.org/lor/80192103-2113.html

14. Moore JR. *Correlation of smoke balance factors with some trends*. 03 Jul 1973. Bates: 2041197322. http://tobaccodocuments.org/pm/2041197322.html

15. Moore JP. *Correlation of pH and share of market performance*. 10 Aug 1973. Bates: 2041197321. http://tobaccodocuments.org/pm/2041197321.html

Trends in mortality from diabetes mellitus in Canada, 1986–2000

Jinfu Hu, Glenn Robbins, Anne-Marie Ugnat, and Chris Waters

Abstract

The purpose of this study was to examine trends in diabetes mellitus (DM) mortality rates in Canada, including analysis at the provincial level, during the period 1986–2000. The study population included Canadians aged 35 and over. Age-standardized mortality rates (ASMRs) were computed. Linear regression was used to calculate the average annual percentage change (AAPC) by age, sex and province. The results showed a substantial increase in DM mortality rates among those aged 35 and over, particularly for men; the AAPC indicated an increase of 2.4% for men and 0.7% for women. When the mortality rates were plotted for three time periods, the rates increased with each successive age group and period for both sexes. Mortality from DM increased significantly in both sexes in Canada between 1986 and 2000, particularly in men

Key Words: *Canada, diabetes mellitus, mortality*

Introduction

Diabetes mellitus (DM) is a growing health problem in Canada. According to data from the National Population Health Survey, the 2-year incidence rate between 1994/95 and 1996/97 was 4.0 cases per 1,000 person-years at risk, but it rose to 6.7 cases between 1998/99 and 2000/01.[1] In 1998, diabetes was the sixth and seventh leading cause of death for Canadian males and females, respectively.[2] The estimated total cost of both diagnosed and undiagnosed diabetes in Canada in 1998 was approximately $5 billion.[3] Diabetes contributes significantly to mortality and reduced life expectancy in elderly subjects.[4] It is estimated that 300 million people worldwide will have developed diabetes by the year 2050.[5]

In this article we update the previous publications on diabetes mortality in Canada,[6,7] and examine trends in mortality rates trends of MD among men and women aged 35 and over in Canada from 1986 to 2000, including analysis at the provincial level.

Methods

Statistics Canada provided data on DM mortality (ICD–9 code 250[8]) for 1986 to1999 and ICD–10 code E10-E14[9] for the year 2000), from annual Canadian mortality files obtained from death certificates. The mortality data used was based on only the "underlying cause of death".[10] The data contained information on age, sex, and province. Population counts were obtained from the Canadian Census.[11] We calculated age-standardized mortality rates using the 1991 Canadian population as the standard.

The study population included Canadians aged 35 and over. We computed age-standardized mortality rates at two levels of geographic aggregation: Canada; provinces (Newfoundland, Prince Edward Island, Nova Scotia, New Brunswick, Quebec, Ontario, Manitoba, Saskatchewan, Alberta, British Columbia) and territories (Yukon Territory, Northwest Territories including Nunavut). We contrasted 5–year age-standardized mortality rates calculated over a 5–year period by sex for the periods 1986–1990, 1991–1995 and 1996–2000.

We used linear regression to calculate average annual percentage change (AAPC) at the national and provincial/territorial levels. The AAPC values from linear regression were determined by fitting a model that assumed a constant rate of change in the age-standardized mortality rates (ASMRs), that is, a linear model applied to the ASMRs after logarithmic transformation. Ninety-five percent confidence intervals (CIs) for age-standardized mortality rates were calculated by province only. We investigated trends by 10–year age groups (35–44, 45–54, 55–64, 65–74, 75–84, 85+), and for the summary grouping (35+). All statistical analyses were performed using SAS software.[12]

Results

Canadian mortality rates from DM increased from 1986 to 2000 among men and women aged 35 and over (Figure 1). The ASMRs increased gradually, but irregularly, from 36.13 per 100,000 in 1986 to 48.39 per 100,000 in 2000 among men, and from 29.67 per 100,000 to 33.28 per 100, 000 among women.

For the entire period of 1986–2000, the overall ASMRs from DM aged 35 and over in Canada were 41.17 per 100,000 among men and 30.30 per 100,000 among women (Table 1). The highest provincial rates during this period occurred in Newfoundland in both sexes (54.72 per 100,000 in men and 53.41 per 100,000 in women). The average annual percentage change (AAPC) indicated a significant increase of 2.4% for men and 0.7% for

Author References

Jinfu Hu, Glenn Robbins, Anne-Marie Ugnat, Chris Waters, Surveillance & Risk Assessment Division, Centre for Chronic Disease Prevention and Control, Public Health Agency of Canada, Ottawa, Ontario, Canada

Correspondence: Jinfu Hu, Surveillance & Risk Assessment Division, Centre for Chronic Disease Prevention and Control, Public Health Agency of Canada, 120 Colonnade Road, AL: 6702A Ottawa, Ontario, Canada K1A 0K9; Fax: (613) 941-2057; E-Mail: jinfu_hu@phac-aspc.gc.ca

TABLE 1
Average annual percentage change (AAPC) in age-standardized mortality rates of DM by sex and province, ages 35 and over, Canada, 1986–2000

	Men			Women		
Province/Terrirory	Mortality rate per 100,000	95% confidence interval	AAPC	Mortality rate per 100,000	95% confidence interval	AAPC
Newfoundland	54.72	(51.08–58.62)	3.30**	53.41	(50.31–56.70)	1.13
Prince Edward Island	31.15	(26.43–36.73)	4.53	23.01	(19.61–27.01)	-4.56
Nova Scotia	37.30	(35.16–39.57)	0.77	28.92	(27.35–30.58)	-0.32
New Brunswick	44.73	(42.09–47.55)	4.18**	38.57	(36.50–40.75)	1.04
Quebec	45.70	(44.74–46.67)	0.49	33.75	(33.10–34.41)	-1.31**
Ontario	44.58	(43.84–45.33)	3.57**	31.50	(30.99–32.02)	2.17**
Manitoba	39.70	(37.75–41.75)	3.04**	29.63	(28.18–31.14)	3.75**
Saskatchewan	35.86	(34.01–37.81)	4.28**	25.98	(24.60–27.45)	1.70*
Alberta	31.08	(29.75–32.46)	2.41**	23.13	(22.16–24.15)	0.07
British Columbia	31.63	(30.63–32.67)	2.48**	22.43	(21.71–23.18)	0.20
Yukon Territory	26.87	(14.71–49.10)	-1.62	33.76	(19.35–58.90)	4.55
Northwest Territories	9.44	(4.28–20.84)	-4.31	16.90	(8.98–31.80)	5.05
Canada	**41.17**	**(40.73–41.61)**	**2.44****	**30.30**	**(29.99–30.61)**	**0.71***

* Significant at p = 0.05
** Significant at p = 0.01

women in Canada between 1986 and 2000; on average, the significant increases in rates were greater than 3.0% per year among men in Newfoundland, New Brunswick, Ontario, Manitoba and Saskatchewan, and were greater than 2.0% per year among women in Ontario and Manitoba. We observed a significant provincial decline in DM mortality rates only among women in Quebec.

We plotted the DM mortality rates by 10–year age group for three 5–year periods within the complete study period (Figure 2). For both sexes, the rates increased with each successive age group. Mortality from DM increased sharply for the 75 and over age groups; the highest rates were observed among those aged 85 or over. The rates also increased with each consecutive time period. For each age group, the highest ASMRs occurred in the most recent time period (1996–2000) in both sexes.

Table 2 presents the ASMRs and AAPC by sex and age for the entire study period (1986–2000). The ASMRs of DM increased with each age group in both sexes; the highest rates appeared among those aged 85 or older: 366.98 per 100,000 for men and 317.39 per 100,000 for women. The AAPC showed a significant increase of 2.0% or more for men aged 55 and over, and of 1.2% and 1.8% for women aged 55–64 and 85 or over, respectively.

Discussion

Our results show that mortality rates from DM in Canada have gradually increased for those over age 35 in both sexes between 1986 and 2000. On average, the annual increase was 2.4% for men and 0.7% for

FIGURE 1
Age-standardized mortality rates (per 100,000) of DM by sex, aged 35 and over, Canada, 1986–2000

TABLE 2
Average annual percentage change (AAPC) in age-standardized mortality rates of DM by sex and age, Canada, 1986–2000

	Men		Women	
Age groups	Mortality rate per 100,000	AAPC	Mortality rate per 100,000	AAPC
35–44	2.56	1.51	1.32	1.56
45–54	7.88	1.24	4.26	-0.73
55–64	26.52	2.74**	17.00	1.21*
65–74	76.58	2.90**	52.42	0.58
75–84	187.82	2.00**	145.39	0.06
85+	366.98	2.85**	317.39	1.84**
35+	41.17	2.44**	30.30	0.71*

* Significant at p = 0.05
** Significant at p = 0.01

women. Rates increased significantly among men aged 55 and over and among women aged 55–64 and 85 or over.

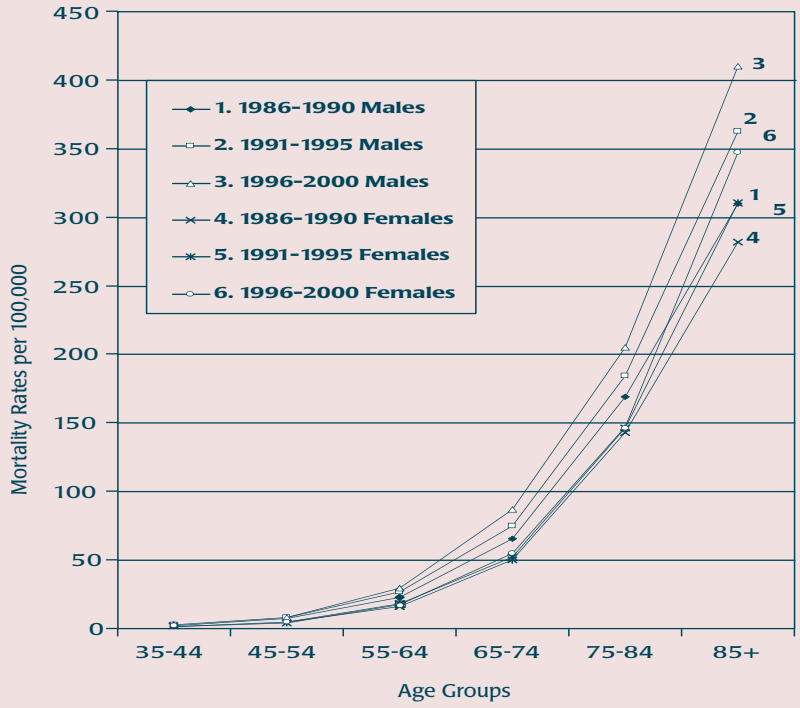
In Missouri, the rate of reported diabetes-related deaths rose by about 3% per year from 1989 to 1994.[13] An increased trend in diabetes mortality was also seen in South Carolina.[14] The diabetes mortality rate for the United States was 13.3 per 100,000 population[15] in 1995 and 25.4 per 100,000 population in 2002.[16] Our findings of increased DM mortality rates are also consistent with those of a previous Canadian study.[6,7]

Increased DM mortality may be due to the increased incidence and prevalence of DM, which is occurring at an alarming rate in both developed and developing countries.[17] In the United States, the prevalence of diagnosed diabetes increased to 4.8% in 200 from 2.9% in 1990.[18] The rapidly growing prevalence of diabetes in Asia heralds a large increase in the incidence of diabetes-related deaths in the coming decades.[19] The population incidence and prevalence of diabetes have also increased in Canada.[1] The age-adjusted prevalence of diabetes rose by 49% between 1986 and 1991 in adults aged 25 or over in Manitoba[20] and increased from 4.7% in 1995 to 6.2% in 1999 in adults aged 20 and over in Ontario.[21] In addition, a large proportion of the population has undetected diabetes.[22,23] For these reasons, the reported prevalence of diabetes is likely a low estimate.

The increase in the prevalence of DM might have different explanations. The prevalence of DM has increased over time while its incidence remains relatively stable in Ontario and Manitoba,[20,21] which might be explained by the fact that people are living longer after diagnosis with DM.[21] We note that the prevalence of MD for all populations in the US was 3.3% in 1996 and 4.1% in 1997.[18] The new American Diabetes Association Diagnostic Criteria,[24] first used in clinical practice in 1997, may have resulted in the detection of a greater proportion of people with undiagnosed diabetes. An increased prevalence of overweight people and obesity may play an important role in the development of DM.

Obesity is known to be a strong risk factor for type 2 diabetes. A strong relationship between obesity and the development of DM has been found in both sexes.[25–27] Most patients with early-onset type 2 diabetes in Mexico are obese or overweight.[28] At present, the prevalence of overweight people and obesity is increasing worldwide,[29] possibly resulting in increased DM incidence and mortality. For Canadian adults aged 20 to 69, the prevalence of overweight people and obesity gradually increased from 47.0% in 1970–1972, to 55.6% in 1978–1979, to 58.1% in 1986–1992 (similar trends were in all age groups) among men. The prevalence of overweight people and obesity was 33.9%, 42.3% and 40.6% among women over the same time period.[30] The prevalence of overweight people and obesity is growing at a faster rate

FIGURE 2
Age-standardized mortality rates (per 100,000) of DM by age, sex and 5–year calendar period, Canada



among men than among women, which may result in rapidly rising mortality rates of DM in men, but not in women. It is possible that obesity heightens the risk of overall mortality and circulatory disease mortality, and even more substantially increases the risk of diabetes mortality.[31]

Our results reveal differences in the ASMRs of DM by geographic region (i.e., province/territory and county), sex and age. These differences may be associated with environmental risk factors and socio-economic status. Sex-related dissimilarities in risk factors were observed in the Finnmark Study[32] and in Uganda.[27] The MONICA Augsburg study also found sex-related differences that seem to be a factor in disease development.[33] High rates of DM prevalence are strongly correlated with indicators of low socio-economic status, poor environmental quality and lifestyle.[34,35] In another Finnish study, a clear socio-economic gradient in mortality emerged in every age group of people with diabetes.[36]

Many studies report excess mortality related to DM. Mortality among individuals with type 2 diabetes has been 1.4 to 3.7 times higher than the rate among people without diabetes,[37] in particular in deaths from circulatory disease.[38–41] Patients with DM have an increased atherogenic cardiovascular risk profile,[42] which might explain the sharp increase in DM mortality rates at ages 75 or over that we found among both sexes in our study.

Type 2 diabetes is one of the lifestyle-related diseases. During a 16–year follow-up a study in the US, researchers found that the majority of cases of type 2 diabetes could have been prevented by the adoption of a healthier lifestyle.[26] Smoking cessation is of utmost importance in facilitating glycemic control and limiting the development of diabetic complications,[43] and may significantly decrease the mortality from DM.[38] Physical activity has been significantly associated with decreased prevalence of type 2 diabetes and related comorbidity,[44] as well as with reduced risk of cardiovascular disease, cardiovascular death and total mortality among men with type 2 diabetes.[45,46] More physical activity and a less sedentary lifestyle are important in preventing obesity and diabetes[44,46–49] and in reducing the risk of cardiovascular events among both sexes.[46,48]

The data in our study have several limitations. One is that the data included all types of diabetes and did not permit the distinction between patients with type 1 and type 2. Although type 1 diabetes most commonly occurs in childhood, it can occur at any age,[50] and could not be ruled out in our study, which focussed on deaths among those aged 35 or older. However, other research reports that type 2 DM currently accounts for about 90–95% of all cases.[50,51] Another limitation is that the mortality data we used here were the underlying cause of death, resulting in significantly under-represented mortality rates. DM is known to be under- reported on death certificates as an underlying or contributing cause of death.[52] In this study, we used data from Statistics Canada's national annual mortality files, and our results on DM mortality may be underestimated. A final limitation is that the data for 2000 were coded according to the ICD–10, unlike the data for the preceding years. However, only a small difference between ICD–9 and ICD–10 codes was seen for diabetes; a preliminary estimate from the US indicates less than a 1% increase due to ICD–10 coding of diabetes.[53]

In conclusion, we observed a substantial increase in mortality from DM among Canadians aged 35 and over between 1986 and 2000, particularly in men. Our results may provide a scientific basis for efforts to reduce mortality from DM among target populations in Canada.

Acknowledgements

The authors gratefully acknowledge Statistics Canada for providing the annual Canadian mortality data to Health Canada from the Canadian Vital Statistics database. The cooperation of the provincial and territorial vital statistics registries, which supply the data to Statistics Canada, is also gratefully acknowledged.

References

1. Millar WJ, Young TK. Tracking diabetes: prevalence, incidence and risk factors. *Health Reports* 2003;14:35–47 (Statistics Canada, Catalogue 82–003).
2. Statistics Canada. *Leading causes of death at different ages, Canada, 1998*. (Catalogue 84F0503XPB). Ottawa: Health Statistics Division, 2001.
3. Dawson KG, Gomes D, Gerstein H, Blanchard JF, Kahler KH. The economic cost of diabetes in Canada, 1998. *Diabetes Care* 2002; 25:1303–1307.
4. Sinclair AJ, Robert IE, Croxson SC. Mortality in older people with diabetes mellitus. *Diabet Med* 1997;14:639–47.
5. Smith RJ. Diabetes mellitus; a 2003 perspective. *Med Health Rhode Island* 2003;86: 92–5.
6. Health Canada. *Diabetes in Canada*. Ottawa: Diabetes Division, 1999.
7. Health Canada. *Diabetes in Canada*. Second edition. Ottawa: Center for Chronic Disease Prevention and Control, Population and Public Health Branch, 2002.
8. World Health Organization. *International classification of diseases, 1975 revision*. Geneva: World Health Organization, 1977.
9. World Health Organization. *International classification of diseases and related health problems, tenth revision*. Geneva: World Health Organization, 1992.
10. Statistics Canada. *Causes of death* (Catalogue 84F208XPB). Ottawa: Health Statistics Division, 1999.
11. Statistics Canada. Census Division Populations. *Annual demographic statistics 2002* (Catalogue 91–213–XPB). Ottawa: Demography Division, 2003.
12. SAS. *SAS System for Windows Release 8.1*. Carey, NC: SAS Institute Inc., 1999–2000.
13. Sprick AN, Simoes EJ, McKeage CB, Chang JC. Diabetes-related deaths in Missouri 1989–1994. *Mo Med* 1998; 95:21–25.
14. Lackland DT, Zheng D, Jones PJ, et al. Epidemiology of diabetes in South Carolina. *JSC Med Assoc* 1998;94:473–77.
15. Anderson RN, Kochanek KD, Murphy SL. *Report of final mortality statistics*, 1995. Monthly vital statistics report; Vol 45 no 11, supp 2. National Center for Health Statistics, 1997.

16. Kochanek KD, Smith BL. *Deaths: Preliminary Data for 2002*. National vital statistics reports. Vol 52 no 13. National Center for Health Statistics. National vital statistics reports. 2004.

17. Williams R, Airey M. Epidemiology and public health consequences of diabetes. *Curr Med Res Opin* 2002;18 Suppl 1:s1–12.

18. Centers for disease Control and Prevention. National Center for Chronic Disease Prevention and Health Promotion. "Prevalence of diabetes: crude and age-adjusted prevalence of diagnosed diabetes per 100 population, by age, United States, 1980-2002." *Data & Trends: Diabetes Surveillance System*. 14 May 2004.

19. Woodward M, Zhang X, Barzi F, et al. The effects of diabetes on the risk of major cardiovascular disease and death in the Asia-Pacific region. *Diabetes Care* 2003;26: 360–3A6.

20. Blanchard JF, Ludwing S, Wajda A, et al. Incidence and prevalence of diabetes in Manitoba, 1986–1991. *Diabetes Care* 1996: 19:807–11.

21. Hux J, Booth G, Laupacis. *Diabetes in Ontario, an ices practice atlas, module 1*. Canadian Diabetes Association, 2002.

22. Lundberg V, Stegmayr B, Asplund K, Eliasson M, Huhtasaari F. Diabetes as a risk factor for myocardial infarction: population and gender perspectives. *J Intern Med* 1997; 241:485–92.

23. Tuomilehto J, Korhoonen HJ, Kartovaara L, et al. Prevalence of diabetes mellitus and impaired glucose tolerance in the middle-aged population of three areas in Finland. *Int J Epidemiol* 1991;20:1010–17.

24. Harris MI, Eastman RC, Cowie CC, Flegal KM, Eberhard MS. Comparsion of diabetes diagnostic categories in the U.S. population according to the 1997 American Diabetes Association and 1980–1985 World Health Organization diagnostic criteria. *Diabetes Care* 1997;20:1859–62.

25. Choi BCK, Shi S. Risk factors for diabetes mellitus by age and sex: results of the National Population Health Survey. *Diabetologia* 2001;44:1221–31.

26. Hu FB, Manson JE, Stampfer MJ, et al. Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. *N Engl J Med* 2001;345: 790–97.

27. Lasky D, Becerra E, Boto W, Otim M, Ntambi J. Obesity and gender differences in the risk of type 2 diabetes mellitus in Uganda. Nutrition 2002;18:417–21.

28. Aguilar-Salinas CA, Rojas R, Gomerz-perez FJ, et al. Prevalence and characteristics of early-onset type 2 diabetes in Mexico. *Am J Med* 2002;113:569–74.

29. World Health Organization. *Obesity: preventing and managing the global epidemic*. Report of a WHO Consultation. Geneva: World Health Organization, 2000.

30. Raine KD. *Overweight and obesity in Canada: a population health perspective*. Canadian Institute for Health Information, Ottawa, 2004.

31. Rogers RG, Hummer RA, Krueger PM. The effect of obesity on overall, circulatory disease and diabetes-specific mortality. *J Biosoc Sci* 2003;35:107–29.

32. Njolstad I, Arnesen E, Lund-Larsen PG. Sex differences in risk factors for clinical diabetes mellitus in a general population: a 12–year follow-up of the Finnmark Study. *Am J Epidemiol* 1998;147:49–58.

33. Meisinger C, Thorand B, Schneider A, et al. Sex differences in risk factors for incident type 2 diabetes mellitus: the MONICA Augsburg cohort study. Arch Intern Med 2002;162:82–9.

34. Green C, Hoppa RD, Young TK, Blanchard JF. Geographic analysis of diabetes prevalence in an urban area. *Soc Sci Med* 2003; 57:551–60.

35. Brancati FL, Whelton PK, Kuller LH, Klag MJ. Diabetes mellitus, race, and socio-economic status. A population-based study. *Ann Epidemiol* 1996;6:67–73.

36. Forssas E, Keskimaki I, Reunanen A, Koskinen S. Widening socio-economic mortality disparity among diabetic people in Finland. *Eur J Public Health* 2003;13:38–43.

37. Muller WA. Diabetes mellitus long-time survival. *J Insur Med* 1998;30:17–27.

38. Koskinen SV, Reunanen AR, Martelin TP, Valkonen T. Mortality in a large population-based cohort of patients with drug-treated diabetes mellitus. *Am J Public Health* 1998;85:765–70.

39. Hu FB, Stampfer MJ, Solomon CG, et al. The impact of diabetes mellitus on mortality from all causes and coronary heart disease in women: 20 years of follow-up. *Arch Intern Med* 2001;161:1717–23.

40. Bloomgarden ZT. Cardiovascular disease and diabetes. Diabetes Care 2003;26: 230–7.

41. Serrano-Rios M. Epidemiology of cardiovascular disease in type 2 diabetes. *Int J Clin Pract* 2001;121 Suppl: 4–7.

42. Hanefeld M, Fischer S, Julius U, et al. Risk factors for myocardial infarction and death in newly detected NIDDM: the Diabetes Intervention Study, 11–year follow-up. *Diabetologia* 1996;39:1577–83.

43. Eliasson B. Cigarette smoking and diabetes. *Prog Cardiovasc Dis* 2003;45:405–13.

44. Wang L, Yamaguchi T, Yoshimine T, et al. A case-control study of risk factors for development of type 2 diabetes: emphasis on physical activity. *J Epidemiol* 2002;12:424–30.

45. Tanasescu M, Leitzmann MF, Rimm EB, Hu FB. Physical activity in relation to cardiovascular disease and total mortality among men with diabetes. *Circulation* 2003;107: 2435–39.

46. Batty GD, Shipley MJ, Marmot M, Simith GD. Physical activity and cause-specific mortality in men with type 2 diabetes/impaired glucose tolerance: evidence from the Whitehall study. *Diabet Med* 2002;19:580–8.

47. Hu FB, Li TY, Colditz GA, et al. Television watching and other sedentary behaviors in relation to risk of obesity and type 2 diabetes mellitus in women. *JAMA* 2003;289: 1785–91.

48. Hu FB, Stampfer MJ, Solomon CG, et al. Physical activity and risk for cardiovascular events in diabetic women. *Ann Intern Med* 2001;134:96–105.

49. Hu FB, Leitzmann MF, Stampfer MJ, et al. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch Intern Med* 2001;161:1542–8.

50. Schuster DP, Duvuuri V. Diabetes mellitus. Clin Podiatr Med Surg 2002;19:79–107.

51. Rock M. Sweet blood and social suffering: rethinking cause-effect relationship in diabetes, distress, and duress. *Med Anthropol* 2003;22:131–74.

52. Fuller JH, Flford J, Goldblatt P, Adelstein AM. Diabetes mortality: new light on an underestimated public problem. *Int J Epidemiol* 2000;29:140–8.

53. Anderson RN, Minino AM, Hoyert DL, et al. Comparability of cause of death between ICD–9 and ICD–10: preliminary estimates. *Natl Vital Stat Rep* 2001;49:1–32.

Status Report

The Canadian Incidence Study of Reported Child Abuse and Neglect (CIS)

Ambika Dewan and Lil Tonmyr

A significant number of Canadian children and youth are victims of abuse and neglect, yet there is little national-level information available on this problem. The Injury and Child Maltreatment Section of the Centre for Healthy Human Development and its many partners across Canada are contributing to our knowledge and understanding of child maltreatment through the Canadian Incidence Study of Reported Child Abuse and Neglect (CIS). The first cycle of data collection for the CIS was in 1998, with a first report issued in 2001. The second cycle occurred in late 2003. Many researchers have used CIS data, in addition to information reported by Public Health Agency of Canada, to examine various aspects of child maltreatment in Canada.

Overview of the CIS

The CIS is the first nationwide study to examine the incidence of reported child maltreatment and the characteristics of children and families investigated by Canadian child welfare services. Information is gathered from participating provincial and territorial child welfare agencies across Canada. The CIS addresses the four principal forms of maltreatment: physical abuse, sexual abuse, neglect and emotional maltreatment.

The CIS is intended to follow the health surveillance cycle, which involves three stages: data collection/acquisition, data analysis, the broad dissemination of the information for action and a return to data collection. The CIS operates on a five-year periodic data collection and reporting cycle. CIS–1998 data collection was conducted in 1998[1] and data collection for CIS–2003 occurred in the fall/winter of 2003. The third cycle of CIS data collection is planned for 2008. The ongoing and cyclical nature of the CIS surveillance system generates data that will continue to build knowledge by tracking trends and forming an evidence base for policy development and evolution of practice. CIS data are also used for detailed research. The first cycle of this surveillance system has already provided valuable information on not only the occurrence of child maltreatment among investigated cases, but also on the characteristics of maltreated children, their families, and their circumstances in the community.

CIS–1998 combined its core study with four provincially supported studies at additional sites in Newfoundland, British Columbia, Quebec and Ontario. A larger sample in these jurisdictions allowed the calculation of independent provincial incidence estimates. Prince Edward Island, the Northwest Territories, Ontario and Alberta are providing extra resources for oversampling in CIS–2003.

Partners

Public Health Agency of Canada carries out the CIS in collaboration with provincial/territorial governments, participating child welfare agencies and a dedicated study team. This is a team of researchers from regional academic centres at the University of Calgary, Memorial University, Université Laval, and the University of Toronto under the leadership of Dr. Nico Trocmé. The study is guided by a national advisory committee in which expertise is drawn from many fields including public health, child advocacy, child welfare, pediatrics, children's mental health, social services and criminal justice.

Present work

Data collection for CIS–2003 is now complete, with the data verified for inconsistencies and entered into a database. The study team will soon enter the data analysis and reporting phase of the study.

The final report for CIS–2003 will be available in October 2005. This report will be the second national report on child maltreatment in Canada and will provide the opportunity to compare child abuse and neglect in Canada at two distinct points in time. When released, copies of the report will be available from the National Clearinghouse on Family Violence by calling 1 800 267-1291 or by e-mailing ncfv-cnivf@hc-sc.gc.ca.

While awaiting the results of CIS–2003, research relating to CIS–1998 is being conducted on many facets of child maltreatment. For example, a recent article[2] presented a comparative analysis of Aboriginal and non-Aboriginal families to identify important differences in terms of numbers of child maltreatment investigations, types of maltreatment, the placement of children in out-of-home-care, and socioeconomic factors. Another article[3] discussed the nature and severity of physical harm caused by abuse and neglect based on reports of child

Author References

Ambika Dewan, Lil Tonmyr, Injury and Child Maltreatment Section, Centre for Healthy Human Development, Health Surveillance and Epidemiology Division, Public Health Agency of Canada, Ottawa, Ontario, Canada

Correspondence: Lil Tonmyr, Injury and Child Maltreatment Section, Centre for Healthy Human Development, Health Surveillance and Epidemiology Division, Public Health Agency of Canada, AL: 1910C, Tunney's Pasture, Ottawa, Ontario, Canada K1A 0L2; Fax: (613) 941-9927; E-mail: Lil_Tonmyr@phac-aspc.gc.ca

maltreatment documented in the CIS. Another research article used CIS data to compare children with developmental delays to those with no developmental delays, in terms of type and severity of abuse and parental characteristics such as income and education.[4]

Important findings from CIS–1998 are also discussed in the latest issue of the Health Policy Research Bulletin, HPRB released in September, 2005. The Health Policy Research Bulletin can be found at www.hc-sc.gc.ca/arad-draa. Copies of CIS-1998 can be ordered from the National Clearinghouse on Family Violence by telephone at 1-800-267-1291 or by e-mail to ncfv-cnivf @hc-sc.gc.ca.

References:

1. Phaneuf G, Tonmyr L. New research initiatives from the Child Maltreatment Division. *Chron Dis Can* 1998;19(2):71–2.

2. Blackstock C, Trocmé N, Bennett M. Child maltreatment investigations among Aboriginal and non-Aboriginal families in Canada. *Violence Against Women* 2004; 10(8):901–16.

3. Trocmé N, MacMillan H, Fallon B, De Marco R. Nature and severity of physical harm caused by child abuse and neglect: results from the Canadian Incidence Study. *CMAJ* 2003;169(9):911–15.

4. Schormans AF, Brown I. An investigation into the characteristics of the maltreatment of children with developmental delays and the alleged perpetrators of this maltreatment. *Journal of Developmental Disabilities* 2002; 9(1):1–19.

Calendar of Events

21– 24 August, 2005 **Nashville, Tennessee, USA**	21st International Conference on Pharmacoepidemiology & Therapeutic Risk Management	Tel.: (301) 718-6500 Fax: (301) 656-0989 E-mail: ispe@paimgmt.com < http://www.pharmacoepi.org/meetings/index.cf >
21– 25 August, 2005 **Bangkok, Thailand**	XVII International Epidemiological Association World Congress on Epidemiology	< http://wce2005.org/index.htm >
18–21 September, 2005 **Ottawa, Ontario, Canada**	Canadian Public Health Association 96th Annual Conference Mapping the Future of Public Health: People, Places and Policies	E-mail: conference @cpha.ca < http://www.cpha.ca/english/conf/96thAnl/96conf.htm >
25–29 September, 2005 **Pretoria, South Africa**	17th Conference of the International Society for Environmental Epidemiology	ISEE 2005 Conference Organizing Committee School of Health Systems and Public Health University of Pretoria PO Box 667 Pretoria 0001 South Africa Tel.: +27 31 266 2384 Fax: +27 31 266 2380 E-mail: confcall@yebo.co.za < http://www.isee2005.co.ze >
19–20 October, 2005 **Edmonton, Alberta, Canada**	9th Annual CDA/CSEM Professional Conference and Annual Meetings	< http://www.diabetes.ca//Section_Professionals/profconf2005.asp >
23–24 October, 2005 **Salt Lake City, UT, USA**	14th Annual Meeting International Genetic Epidemiology Society	< http://www.biostat.wustl.edu/ ~ genetics/iges/meetings.html >
23–26 October, 2005 **Vancouver, British Columbia**	1st International Cancer Control Congress	< http://www.cancercontrol2005.com >
5–9 November, 2005 **New Orleans, LA, USA**	133rd Annual Meeting: Evidence Based Policy and Practice American Public Health Association	E-mail: diane.lentini@apha.org < http://www.apha.org/meetings/future_past.htm >
6–8 May, 2006 **Montréal, Quebec, Canada**	Reasons for Hope 2006 – CBCRA's 4th Scientific Conference Canadian Breast Cancer Research Alliance	Susan Wall Coordinator, Conferences and Meetings Tel.: (416) 596-6598 x 313 E-mail: swall@cbcf.org < http://www.breast.cancer.ca/reasons_for_hope_conferences/Default.asp?language = English >

Indexes for Volume 25, 2004

***Note: In the Spring 2004 issue of CDIC, page numbering was inadvertently re-started at 1.**

Volume 25 Contents

No 1, 2004

No 2, 2004

No 3/4, 2004

INJURIES

Rates of carpal tunnel syndrome, epicondylitis, and rotator cuff claims in Ontario workers during 1997. 25(2):32-38.

OCCUPATIONAL HEALTH

Occupational exposure to chemical and petrochemical industries and bladder cancer risk in four western Canadian provinces. 25(2):7-15.

POPULATION SURVEILLANCE

A model for non-communicable disease surveillance in Canada: The Prairie Pilot Diabetes Surveillance System. 25(1):7-12.

Computer assisted telephone interviewing (CATI) for health surveys in public health surveillance: Methodological issues and challenges ahead. 25(2):21-27.

Methods for estimating the labour force ensured by the Ontario Workplace Safety & Insurance Board: 1990-2000. 25(3/4):127-37.

TOBACCO ISSUES

Youth attitudes towards tobacco control: A preliminary assessment. 25(3/4):97-100.

TRAUMA

Rates and external causes of blunt head trauma in Ontario: Analysis and review of Ontario Trauma Registry datasets. 25(1):32-41.

Rates of claims for cumulative trauma disorder of the upper extremity in Ontario workers during 1997. 25(1):22-31.

WOMEN'S HEALTH

Do work-related breast cancer risks in pre-menopausal women depend on family history? 25(3/4):147-51.

Volume 25 Author Index

New Editor-in-Chief

As past Editor-in-Chief of *Chronic Diseases in Canada*, I extend my sincere apologies to the authors, subscribers and everyone else who has been affected by the recent unprecedented delay in the processing and review of manuscripts and in the publication of this issue.

The move from Health Canada to the new Public Health Agency of Canada, a further move within the Agency, and personnel changes all caused disruptions throughout the system.

I am pleased to announce that David Carle-Ellis has agreed to take on the position of Acting Editor-in-Chief of *Chronic Diseases in Canada.* The transition will be gradual; we will both be available to ensure the continuity of the journal.

CDIC: Information for Authors

Chronic Diseases in Canada (CDIC) is a peer-reviewed, quarterly scientific journal focussing on the prevention and control of non-communicable diseases and injuries in Canada. This may include research from such fields as epidemiology, public/community health, biostatistics, behavioural sciences and health services. CDIC endeavours to foster communication on chronic diseases and injuries among public health practitioners, epidemiologists and researchers, health policy planners and health educators. Submissions are selected based on scientific quality, public health relevance, clarity, conciseness and technical accuracy. Although CDIC is a publication of the Public Health Agency of Canada, contributions are welcomed from both the public and private sectors. Authors retain responsibility for the contents of their papers, and opinions expressed are not necessarily those of the CDIC Editorial Committee or of the Public Health Agency of Canada.

Feature Articles

Regular Feature Articles: Maximum 4,000 words for main text body (excluding abstract, tables, figures, references) in the form of original research, surveillance reports, meta-analyses, methodological papers, literature reviews or commentaries.

Short Reports: Maximum 1,200 words (as above).

Status Reports: Describe ongoing national programs, studies or information systems at the Public Health Agency of Canada (maximum 3,000 words).

Workshop/Conference Reports: Summarize workshops, etc. organized or sponsored by the Agency (maximum 3,000 words).

Cross-Canada Forum: For authors outside of the Public Health Agency of Canada to exchange information from research or surveillance findings, programs under development or program evaluations (maximum 3,000 words).

Additional Article Types

Letters to the Editor: Comments on articles recently published in CDIC will be considered for publication (maximum 500 words).

Book/Software Reviews: Usually solicited by the editors (500–1,300 words), but requests to review are welcomed.

Submitting Manuscripts

Submit manuscripts to the Editor-in-Chief, *Chronic Diseases in Canada*, Public Health Agency of Canada 130 Colonnade Road, CDIC Address Locator: 6501G, Ottawa, Ontario K1A 0K9, e-mail: cdic-mcc@phac-aspc.gc.ca.

Since CDIC adheres in general (section on illustrations not applicable) to the "**Uniform Requirements for Manuscripts Submitted to Biomedical Journals**" as approved by the International Committee of Medical Journal Editors, authors should refer to this document for complete details before submitting a manuscript to CDIC (see **<www.cma.ca>** or Can Med Assoc J 1997;156(2): 270–7).

Checklist for Submitting Manuscripts

Cover letter: Signed by all authors, stating that all have seen and approved the final manuscript and have met the authorship including a full statement regarding any prior or duplicate publication or submission for publication.

First title page: Concise title; full names of all authors and institutional affiliations; name, postal and e-mail addresses, telephone and fax numbers for corresponding author; separate word counts for abstract and text.

Second title page: Title only; start page numbering here as page 1.

Abstract: Unstructured (one paragraph, no headings), maximum 175 words (100 for short reports); include 3–8 key words (preferably from the Medical Subject Headings (MeSH) of Index Medicus).

Text: Double-spaced, 1 inch (25 mm) margins, 12 point font size.

Acknowledgements: Include disclosure of financial and material support in acknowledgements; if anyone is credited in acknowledgements with substantive scientific contributions, authors should state in cover letter that they have obtained written permission.

References: In "Vancouver style" (consult Uniform Requirements and a recent CDIC issue for examples); numbered in superscript (or within parentheses) in the order cited in text, tables and figures; listing up to 6 authors (first 3 and "et al." if more); without any automatic reference numbering feature used in word processing; any unpublished observations/data or personal communications used (discouraged) to be cited in the text in parentheses (authors responsible for obtaining written permission); authors are responsible for verifying accuracy of references.

Tables and Figures: Each on a separate page and in electronic file(s) separate from the text (not imported into the text body); as self-explanatory and succinct as possible; not duplicating the text, but illuminating and supplementing it; not too numerous; numbered in the order that they are mentioned in the text; explanatory material for tables in footnotes, identified by lower-case superscript letters in alphabetical order; figures limited to graphs or flow charts/templates (no photographs), with software used specified and titles/footnotes on a separate page.

Number of copies: If submitting by mail, one complete copy, including tables and figures; one copy of any related supplementary material, and a copy of the manuscript on diskette. If submitting by e-mail to cdic-mcc@phac-aspc.gc.ca, please fax or mail the covering letter to the address on the inside front cover.

Background photo and
Inset photo: Digital Stock