Catalogue no. 82-003-X
ISSN 1209-1367

Health Reports

Linkage of the nationally representative Canadian Community Health Survey – Nutrition 2004 to routinely collected mortality records

by Mahsa Jessri, Deirdre Hennessy, Anan Bader Eddeen, Carol Bennett, Didier Garriguet, Claudia Sanmartin, and Douglas Manuel

Release date: September 15, 2022

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Linkage of the nationally representative Canadian Community Health Survey – Nutrition 2004 to routinely collected mortality records

by Mahsa Jessri, Deirdre Hennessy, Anan Bader Eddeen, Carol Bennett, Didier Garriguet, Claudia Sanmartin, and Douglas Manuel

DOI: https://www.doi.org/10.25318/82-003-x202200900002-eng

ABSTRACT

Introduction

The Canadian Community Health Survey (CCHS) – Nutrition 2004 (n=35,107; interview dates from January 2004 to January 2005) linked to the Canadian Vital Statistics – Death Database (CVSD) (2011) represents a novel linkage of a population-based, nationally representative nutrition survey with routinely collected mortality records (including date and cause of death). The linkage was done through individual tax data in Canada, and contains longitudinal records for 29,897 Canadians aged 0 years and older—1,753 of whom died—in the 10 provinces of Canada. The median follow-up time was 7.49 years, with 102,953 person-years among males and 114,876 person-years among females (unweighted), and included a special sampling survey weight (for linked data) to account for those who did not agree to share and link their information. The CCHS – Nutrition 2004 linked to CVSD has been used to evaluate associations between lifestyle and sociodemographic characteristics and mortality. Using these data, statistical methods have been developed and tested to control random and systematic measurement errors when evaluating the relationship between different dietary exposures (evaluated using repeated 24-hour dietary recalls) and health outcomes. The linked data are available through Statistics Canada's Research Data Centres.

Keywords

Statistics Canada's Data Linkage Program, National Survey Data, Vital Statistics, Universal Healthcare System

AUTHORS

Masha Jessri (mahsa.jessri@ubc.ca) is with the Food, Nutrition and Health Program and the Centre for Health Services and Policy Research, at the University of British Columbia. Mahsa Jessri, Deirdre Hennessy, Didier Garriguet, Claudia Sanmartin, and Douglas Manuel are with the Health Analysis Division at Statistics Canada, Ottawa, Ontario. Anan Bader Eddeen, Carol Bennett, and Douglas Manuel are with the Institute for Clinical Evaluative Sciences in Ottawa, Ontario. Carol Bennett and Douglas Manuel are also with the Clinical Epidemiology Program at the Ottawa Hospital Research Institute, and Douglas Manuel is also with the Department of Family Medicine at the University of Ottawa, and the C.T. Lamont Primary Health Care Research Centre Program at the Bruyère Research Institute in Ottawa, Ontario, Canada.

National health surveys and vital statistics registries are the cornerstones of surveillance, monitoring and policy development in most developed countries.[1] National health surveys collect detailed information on a broad range of health behaviours, health statuses and sociodemographic characteristics from large, representative samples of populations, with the aim of monitoring risk factors and health status at the national level and to inform policy.[2] However, these surveys are cross-sectional, and therefore no information on participants' health status is collected beyond the survey date (via self-report). By contrast, vital statistics registries provide objective measures of mortality at the national population level, despite being limited in providing any information beyond basic demographic characteristics.[1,3,4] Increasingly, vital statistics are being linked to national health surveys,[5-14] thus providing a unique and prospective component to these nationally representative surveys for in-depth assessment of mortality determinants[8-12] including socioeconomic and lifestyle behaviours, such as dietary intake, smoking and physical activity.[15-21] Particularly, national nutrition surveys are not routinely linked with health administrative databases—this results in a lack of evidence on the relationship between health and unhealthy diets at the population level.

Aims

This cohort profile describes the linkage of the Canadian Community Health Survey (CCHS) – Nutrition (2004) to the Canadian Vital Statistics – Death Database (CVSD) (2011). Furthermore, it demonstrates how these data can be used in terms of potential analysis strategies and their strengths and limitations—considerations that could apply to any national nutrition survey worldwide.

In addition, a summary of findings on the associations between all-cause mortality and selected sociodemographic characteristics and lifestyle behaviours are presented. This resource provides the first nationally representative linked nutrition survey in Canada, allowing for a generalizable evaluation of the potential link between dietary intake and mortality.

Data resource description

Canadian Community Health Survey – Nutrition

Statistics Canada has collected cross-sectional, nationally representative CCHS data since 2000 with the objective of providing timely information on health status and determinants, as well as health system utilization. The CCHS is funded through the partnership of Statistics Canada, the Public Health Agency of Canada, Health Canada and the Canadian Institutes for Health Information. Since the inception of the CCHS, it was determined that a focused survey cycle (longer version) dedicated to collecting reliable and timely information on dietary intakes was required to complement the information on other health determinants. The CCHS –Nutrition survey was conducted in 2004 to provide detailed information on usual dietary intakes (in terms of food groups, nutrients, nutritional supplements and eating patterns) and measure the prevalence of household food security.[22-24] The CCHS – Nutrition 2004 (January 2004 to January 2005) was a complex, multistage cross-sectional survey that included 35,107 non-institutionalized Canadians aged 0 years and older from 10 provinces (response rate=76.5%, representing 98% of residents in the provinces).[22-24] Data on sociodemographic characteristics, lifestyle behaviours and selected health conditions were collected using interviewer-administered questionnaires, and weight and height were directly measured by trained interviewers. A modified version of the United States Department of Agriculture (USDA) automated multiple-pass method (AMPM) was used to collect two standardized 24-hour dietary recalls,[25,26] with the second recalls collected approximately 3 to 10 days after the first interview on a different day of the week, for about 30% of the sample. The second dietary recall can be used to assess the day-to-day variation in dietary intakes for the entire sample—including those with only one day of recalls—when using the National Cancer Institute (NCI) method.[27,28] The nutrient composition of foods was analyzed using Health Canada's Canadian Nutrient File.[29] Listing below presents the CCHS – Nutrition 2004 components:

- 24-hr recall (all ages)
- Administration (data sharing) (all ages)
- Alcohol (age ≥ 12 years)
- Children's physical activity (age 6 to 11 years)
- Chronic conditions (all ages)
- Food security (all households)
- Fruit and vegetable consumption (age ≥ 6 months)
- General health (age ≥ 12 years)
- Household and education (all ages)
- Income (all ages)
- Labour force participation (age 15 to 75 years)
- Measured height and weight (age ≥ 2 years)
- Physical activities (age ≥ 12 years)
- Sedentary activities (age 12 to 17 years)
- Self-reported height and weight (subset age ≥ 18 years)
- Smoking (age ≥ 12 years)
- Sociodemographic characteristics (all ages)
- Vitamin and mineral supplements (all ages)
- Vitamin and mineral supplement details (all ages)
- Women's health (age ≥ 9 years)

Table 1
Crude unadjusted rates of all-cause mortality by follow-up year among Canadian Community Health Survey – Nutrition 2004 participants who were linked to the Canadian Vital Statistics – Death Database, 2004/2005 to 2011 (n=29,897 aged 0 years and older)

	All-cause mortality rates	
Follow-up year	Unweighted	Weighted
	percent	percent
Less than and equal to 1	0.37	0.36
More than 1 to 2	0.63	0.55
More than 2 to 3	0.65	0.47
More than 3 to 4	0.79	0.56
More than 4 to 5	0.89	0.63
More than 5 to 6	0.90	0.76
More than 6 to 7	0.90	0.56
More than 7 to 8	0.87	0.64

Sources: Canadian Community Health Survey – Nutrition 2004; Canadian Vital Statistics – Death Database.

More details on the survey are available elsewhere.[22]

The Canadian Vital Statistics – Death Database (formerly the Canadian Mortality Database)

The CVSD is a census of all deaths registered in Canada since 1950. Information on deaths is collected by Statistics Canada from provincial and territorial vital statistics registries. The CVSD includes demographic (name, date of birth, date of death, postal code at the time of death) and medical (cause of death) information. The cause of death is coded using the World Health Organization's International Classification of Diseases, version 10. Deaths occurring between January 1, 2004, and December 31, 2011, were eligible for linkage to the CCHS – Nutrition survey **(Table 1)**. More information regarding the CVSD is available elsewhere.[30]

Data collected

Linkage of the CCHS – Nutrition 2004 was conducted as part of a more comprehensive linkage of the CCHS national health surveys to mortality data.[1] Data were collected under the authority of the *Statistics Act* and were analyzed and accessed at Statistics Canada. Of the 35,107 participants of the CCHS – Nutrition 2004, 89.33% (n=29,897) consented to share their survey information with federal and provincial ministries of health and to link their responses to administrative databases. Special sampling survey weights were created for the linked CCHS – Nutrition 2004 data to adjust for those who did not agree to share and link their information.

The linkage was conducted based on Fellegi-Sunter's theory of probabilistic record linkage using G-Link, the generalized record linkage system developed by Statistics Canada to support large-scale record linkages.[31,32] The following unique identifiers were used in the linkage process: given name, last name, date of birth, sex and postal code. The Historical Tax Summary File (HTSF) was used to assist in record linkage through the provision of additional linkage information (given name, last name, postal code) and in the manual resolution of doubtful links. The HTSF is a compilation of tax return files and represents unique individuals for whom a tax declaration was created in a given year. Internal and external validation was conducted to evaluate the accuracy of the linkage process and the fitness of the data for use in the analysis. Global false positive and false negative rates for the more comprehensive linked data were estimated at 0.04% and 2.43%, respectively, using a weighting function. Further information regarding the linkage process has been described previously.[1]

Data linkage was approved by Statistics Canada's Executive Management Board.[33] The use of the data was governed by the Directive on Record Linkage.[34] Participants' privacy during record linkage and use of the linked files were ensured by Statistics Canada. Access to the unique identifying information (e.g., names) was limited to employees directly involved in linking the databases, and these individuals did not access health information. After record linkage, all identifying information was removed (de-identified), and an analytical file was created for subsequent use and analysis.

The record linkage was conducted by Statistics Canada with funding support from Health Canada, the Ottawa Hospital Research Institute and the Canadian Institutes for Health Research.

Quality assessment

Overall, 5.9% of CCHS – Nutrition respondents (n=1,753) eligible for linkage died and were linked to a mortality record. **Table 1** presents the crude unadjusted mortality rates by follow-up year in the CCHS – Nutrition 2004 data linked with the CVSD. The median follow-up time was 7.49 years, with 102,953 person-years for males and 114,876 person-years for females (unweighted) as of December 2011. Linkage rates were compared across respondents' characteristics (**Table 2**) to assess whether the pattern of linkage rates reflected the differential mortality risk among groups at a higher risk.[35,36] Generally, the linkage rates (weighted) were higher for older adults 75 years and older (39.94% vs. 0.09% for children 12 years and younger) and males (5.06% vs. 3.85% for females), reflecting the expected higher mortality risks in these groups. Participants in Alberta, British Columbia and Ontario were less likely to link to the CVSD than those in other provinces, while participants in Nova Scotia were more likely to be linked. The overall pattern of linkage was in line with the general CCHS

Table 2
Number and percentage of the Canadian Community Health Survey – Nutrition 2004 participants who were linked to the Canadian Vital Statistics – Death Database by sex, province and age group, 2004/2005 to 2011 (n=29,897 aged 0 years and older)

	Eligible CCHS – Nutrition 2004		CCHS – Nutrition 2004 Linked to		Linkage rate	
Variables	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
	number				percent	
Total	29,897	31,030,722	1,753	1,380,179	5.86	…
Sex						
Males	14,146	15,326,295	826	775,073	5.84	5.06
Females	15,751	15,704,427	927	605,106	5.89	3.85
Province						
Newfoundland and Labrador	1,527	512,487	82	24,976	5.37	4.87
Prince Edward Island	1,203	135,973	80	7,028	6.65	5.17
Nova Scotia	1,551	909,560	77	61,501	4.96	6.76
New Brunswick	1,463	729,794	76	41,974	5.19	5.75
Quebec	4,244	7,369,295	199	360,184	4.69	4.89
Ontario	8,984	12,176,330	729	487,244	8.11	4.06
Manitoba	3,623	1,093,465	173	53,212	4.78	4.87
Saskatchewan	1,788	925,094	103	54,308	5.76	5.87
Alberta	2,468	3,107,881	92	124,584	3.73	4.01
British Columbia	3,046	4,070,843	142	165,169	4.66	4.06
Age groups (at time of interview)						
<12	7,381	4,317,565	5	3,879	0.07	0.09
≥12 to <25	7,495	5,611,420	29	29,601	0.39	0.53
≥25 to <34	2,373	3,882,970	9	11,135	0.38	0.29
≥35 to <44	2,261	5,078,373	24	33,277	1.06	0.66
≥45 to <54	2,937	4,955,398	85	111,757	2.89	2.26
≥55 to <64	2,578	3,373,464	185	182,884	7.18	5.42
≥65 to <74	2,255	2,046,551	370	302,687	16.41	14.79
≥75	2,617	1,764,982	1,046	704,959	39.97	39.94

… not applicable

Notes: CCHS: Canadian Community Health Survey; CMDB: Canadian Mortality Database.

Sources: Canadian Community Health Survey – Nutrition 2004; Canadian Vital Statistics – Death Database.

linkages conducted by Statistics Canada.[1] Further external validation was conducted to compare the mortality patterns in the linked data with official mortality statistics.[37] The distribution of death across several variables for the linked CCHS – Nutrition data is presented in **Table 3**. The age distribution in the linked CCHS – Nutrition data showed an increased occurrence of deaths with age, with about 51.08% of these deaths concentrated in those aged 75 years and older (weighted estimate). A similar pattern was observed in the CVSD dataset (60.7%), considering that deaths among the institutionalized population are not captured in the linked CCHS – Nutrition 2004 data since institutional residents were not part of the CCHS survey. Geographic distributions of deaths in the linked CCHS were also similar to those derived from the CVSD, with the majority of deaths occurring in the largest provinces (Ontario and Quebec). As expected, the weighted distributions of deaths resembled those in the CVSD more closely than the unweighted distributions. The distribution of deaths by sex and cause were also similar to those in the CVSD dataset.

Figure 1 and Table 4 presents the age-standardized mortality rates (ASMRs) derived from the linked CCHS – Nutrition 2004 file compared with the official mortality rates. Annual adjusted mortality rates are from the Canadian population in 2001. As expected, annual ASMRs in the linked CCHS – Nutrition 2004 file were lower than those derived from official mortality databases reported by Statistics Canada because linkage does not capture 100% of the population. People living on First Nations reserves, people living on Crown lands, institutional residents and full-time Canadian Forces members were excluded from the CCHS sampling. There is also a risk of healthy respondent bias in which those who respond to surveys tend to be healthier than non-respondents. The selection bias could affect the mortality rate in the years in this cohort but would likely be attenuated over time. The ASMRs were lowest in the first year of follow-up for the CCHS – Nutrition 2004, and generally increased in the following years.

Data resource use

This new data resource has several potential applications, and can be used to evaluate associations between lifestyle and sociodemographic characteristics and mortality. An example of such an application is presented below. The analyses presented in the following section pertain to respondents aged 20 or older from the CCHS – Nutrition 2004 linked to CVSD data (unweighted person-years: 117,323) since the mortality rate in those aged younger than 20 is low; including those aged younger than 20 years means there is not sufficient power for the analysis.

Example: Application of the CCHS – Nutrition 2004 linked to CVSD data

Here a brief example of how the data can be used in the evaluation of the relationship between dietary factors (i.e., sodium, saturated fats and added sugars intakes) and mortality is presented.

Table 3

Unweighted and weighted distribution of mortality among Canadian Community Health Survey – Nutrition 2004 linked to the Canadian Vital Statistics – Death Database participants, by sex, province, age group and causes of death, 2004/2005 to 2011 (n=1,753 aged 0 years and older), for each day of the 24-hour recalls

	Day 1		Day 2	
Variables	**Unweighted**	**Weighted**	**Unweighted**	**Weighted**
		percent		
Sex				
Males	5.84	5.06	6.02	6.64
Females	5.89	3.85	5.64	4.40
Province				
Newfoundland and Labrador	5.37	4.87	4.91	3.65
Prince Edward Island	6.65	5.17	6.50	4.95
Nova Scotia	4.96	6.76	4.56	5.60
New Brunswick	5.19	5.75	5.86	7.11
Quebec	4.69	4.89	5.90	7.02
Ontario	8.11	4.00	6.66	3.27
Manitoba	4.78	4.87	5.38	5.18
Saskatchewan	5.76	5.87	6.73	7.13
Alberta	3.73	4.01	5.10	5.89
British Columbia	4.66	4.06	5.54	4.89
Age groups				
<12	F	F	…	…
≥12 to <25	0.39	0.53	0.50	0.38
≥25 to <34	0.38	0.29	…	…
≥35 to <44	1.06	0.66	…	…
≥45 to <54	2.89	2.26	2.06	1.62
≥55 to <64	7.18	5.42	8.20	6.70
≥65 to <74	16.41	14.79	16.26	13.75
≥75	39.97	39.94	37.67	40.63
Causes of death (ICD-10 codes)				
Diseases of the blood and blood-forming organs and certain disorders involving the immune system (D50–D89)	F	F	…	…
Cancers (C00–D48)	28.92	32.3037	28.36	31.1782
Diseases of the circulatory system (I00–I99)	33.31	29.0921	34.89	29.0083
Diseases of the digestive system (K00–K93)	3.99	3.4434	3.92	3.5724
Endocrine, nutritional and metabolic diseases (E00–E90)	3.82	2.4206	3.36	2.4926
External causes of morbidity and mortality (V01–Y98)	6.50	8.8124	6.90	6.7517
Diseases of the genitourinary system (N00–N99)	2.68	2.1481	2.61	3.0022
Certain infectious and parasitic diseases (A00–B99)	1.83	1.8203	1.68	2.5946
Congenital malformations, deformations and chromosomal abnormalities (Q00–Q99)	…	…	…	…
Mental and behavioural disorders (F00–F99)	4.39	4.2804	3.73	4.8462
Diseases of the musculoskeletal system and connective tissue (M00–M99)	0.68	0.5368	…	…
Diseases of the nervous system (G00–H95)	3.25	3.7509	2.99	2.8285
Diseases of the respiratory system (J00–J99)	9.30	10.1676	9.51	12.0791
Diseases of the skin and subcutaneous tissue (L00–L99)	0.00	0.0000	0.00	0.0000
Symptoms, signs and abnormal clinical and laboratory findings and chromosomal abnormalities (R00–R99)	0.91	0.6121	1.31	1.0942
Other causes	0.00	0.0000	0.00	0.0000

… not applicable

F too unreliable to be published

Note: ICD-10 is the 10th revision of the International Classification of Diseases and Related Health Problems.

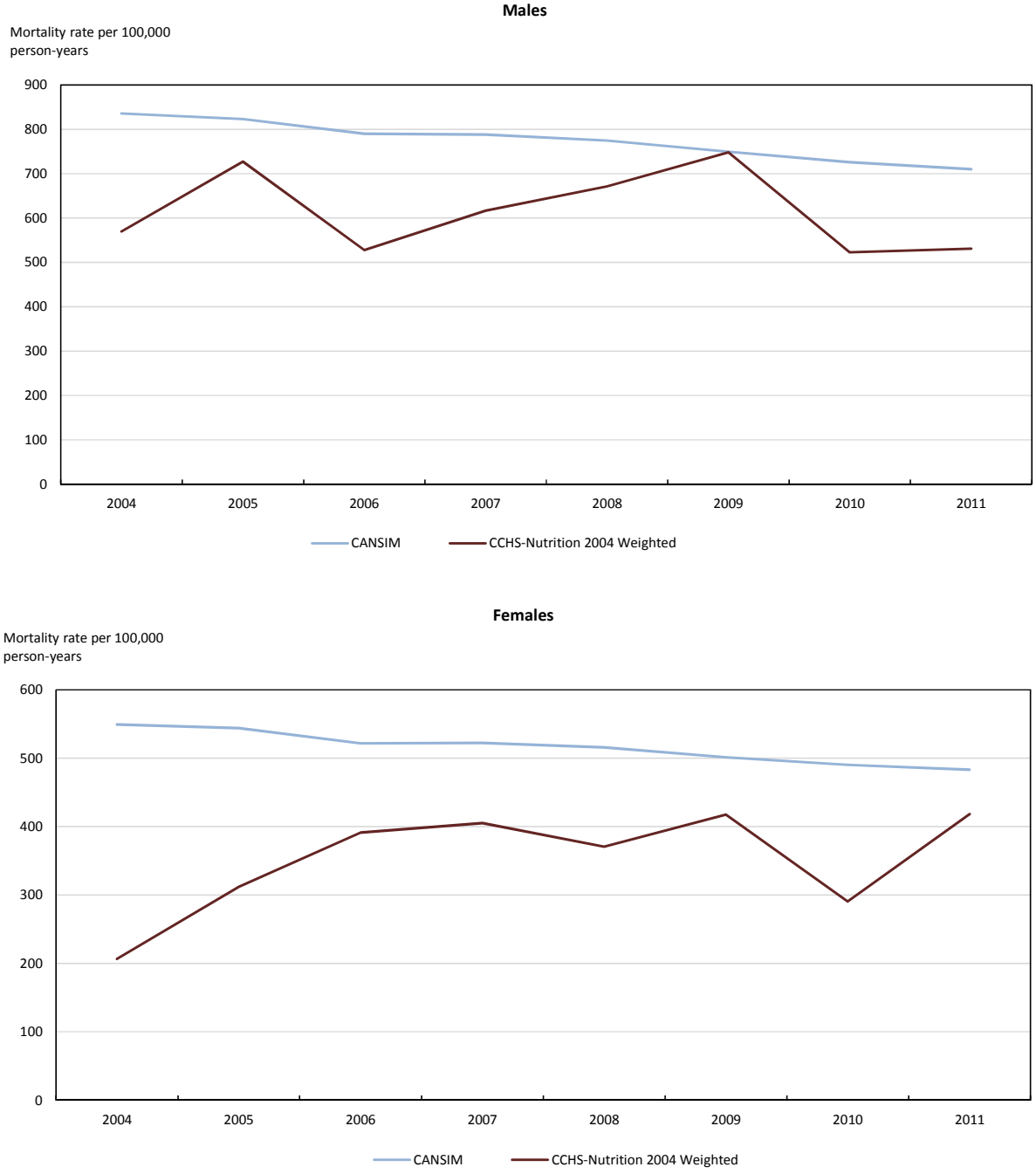
Sources: Canadian Community Health Survey – Nutrition 2004; Canadian Vital Statistics – Death Database.

Data: This analysis included 16,438 participants aged 20 years and older. For analyses of nutrients (sodium, saturated fats and added sugars intakes), 24-hour recalls were used. Invalid 24-hour dietary recalls were removed, as well as records for participants who were pregnant or breastfeeding, leaving us with 16,212 (n=4,901 of whom had a second recall) complete dietary recall data for nutrient-specific analyses (median follow-up: 7.48 years). Outliers were identified and winsorized through observing the model residuals after Box-Cox transformation of exposure (diet) variable, and all analyses accounted for age as well as the nuisance effects including the indicator of sequence of diet recall (first vs. second) and the day of the week when the 24-hour dietary recall was collected (weekend vs. weekday). The distribution of usual intakes for sodium density (grams per 1000 kcal)[38,39] and percentage of energy from saturated fat and added sugars[40] were estimated using the NCI method, which can be used to account for day-to-day variation in intakes when more than one day of intake is available for some of the respondents.[41,42] Since added sugar values are not included in the CCHS food composition database, an estimation method proposed by Brisbois et al. and previously tested by our team[43-47]was used to derive approximate estimates of added sugars.

Variables: Mortality, dietary variables (i.e., sodium, saturated fats and added sugars intakes), and covariates (i.e., age).

Analysis: In the first step of the NCI analysis, univariate likelihood-based models were used to estimate the distribution of intake (simulated population) and to generate intake percentiles. In the next step, regression calibration models were fitted for estimating the association of nutrients of public health concern with all-cause mortality. Age-adjusted Cox proportional hazard ratios (HRs) (95% confidence intervals [CIs]) were calculated for all-cause mortality, with the usual

Figure 1
Age-standardized mortality rates per 100,000 person-years, by sex and survey calendar year, Canadian Community Health Survey – Nutrition linked to Canadian Vital Statistics – Death Database compared with the Canadian Socioeconomic Information Management System, 2004/2005 to 2011 (n=16,438 aged 20 years and older)



Notes: CANSIM: The Canadian Socio-Economic Information Management System ;CCHS: Canadian Community Health Survey.
Source: Canadian Community Health Survey - Nutrition 2004.

Table 4
Age-standardized mortality rates per 100,000 person-years, by sex and survey calendar year, Canadian Community Health Survey – Nutrition linked to the Canadian Vital Statistics – Death Database compared with the Canadian Socioeconomic Information Management System, 2004/2005 to 2011 (n=16, 438 aged 20 years and older)

	Males		Females	
Calendar year	ASMR unweighted	ASMR weighted	ASMR unweighted	ASMR weighted
2004	449.89	648.41	200.56	216.05
2005	642.24	806.99	395.66	367.74
2006	818.57	654.00	572.36	529.48
2007	901.76	966.27	627.42	545.20
2008	1101.53	992.91	659.91	541.17
2009	1073.59	1140.57	842.03	715.08
2010	1243.13	1064.49	816.73	546.25
2011	1101.86	1017.75	826.89	857.31

Note: ASMR = Age-standardized mortality rate.
Sources: Canadian Community Health Survey – Nutrition 2004; Canadian Vital Statistics – Death Database.

intake of nutrients entered as continuous in all models. All 95% CIs for HRs were estimated from coefficient of variations after 500 bootstrap runs. The misreporting status was used to address systematic measurement errors in dietary intakes. Outliers in terms of dietary intakes were identified and replaced with the next closest plausible value to address random measurement errors.

Results: **Table 5** presents the crude weighted mean of the estimated usual intakes of sodium density, percentage of energy from added sugars and saturated fat. Generally, weighted and age-adjusted HRs were revealed increased mortality with each 100-unit increase in the usual intake of sodium density in males (HR: 1.06 95% CIs: 0.95 to 1.17), even though it did not reach statistical significance. Each 5% increase in usual intake distribution of added sugars was associated with a significant 1.32 (1.1 to 1.57) times higher all-cause mortality risk in females, but not in males. As expected, the association of the sodium density, percentage of energy from saturated fats and added sugars with all-cause mortality were weaker when using single 24-hour dietary recalls. Adjustment for usual intake distribution of energy intakes using bivariate likelihood-based models did not change the magnitude and direction of results significantly.

Future work: Under the same funding umbrella used to create this database, a more in-depth and completely adjusted analysis of nutrients (i.e., sodium, added sugars, saturated fats) and dietary patterns in relation to mortality has been completed. These data are published elsewhere.[48]

Considerations: These analyses are not fully adjusted and are for presentation of proper data use only, and therefore should not be interpreted as complete adjusted analyses. Reliance of CCHS – Nutrition 2004 linked to Canadian Vital Statistics – Death Database on self-reported dietary data from 24-hour recalls poses challenges and limitations when evaluating long-term outcomes (e.g., mortality). The use of appropriate methodology for handling random and systematic measurement errors is essential to avoid misleading and conflicting results. Systematic errors can cause the estimate to be biased in any direction.[49] Implementing these methodologies requires specialized training. Overall, the analyses of these data require careful planning and attention to the presence of measurement errors, especially when evaluating long-term outcomes.

Strengths and weaknesses

The newly created CCHS – Nutrition linked to CVSD has several strengths. First, it is a population-based nationally representative nutrition survey linked with mortality, which provides a rich dataset to examine the relationship between dietary intake and health outcomes. It is also straightforward to update the dataset with additional follow-up years as the data becomes available at Statistics Canada. Longitudinal studies provide a powerful resource to investigate the association of dietary intakes with health outcomes. Most longitudinal studies in the field of nutritional science have been cohort studies using food frequency questionnaires for collecting dietary data; these cohorts have a less complex sampling framework because they are not designed to be representative of the population. While the emergence of record linkage techniques has enabled us to link detailed survey information to outcomes, the analyses of nutritional data in this context warrant careful planning and attention to the presence of random and systematic errors.[50] It is necessary to use specific methodologies, such as those described in this paper, to ensure robust estimates, especially when evaluating long-term outcomes (e.g., mortality). There are also some provinces (e.g., Alberta) that were less likely to have CCHS – Nutrition 2004 respondents linked with the CVSD. Although it is unclear why this may be, this discrepancy should be a caution when using the data.

With regard to the example presented, there are limitations and challenges to using nationally representative linked nutrition surveys with 24-hour recalls as the main method of data collection (e.g., CCHS – Nutrition 2004 linked with CVSD). These challenges and limitations include the time-consuming, computationally intensive analyses and the need for proper adjustment for measurement errors in analyses to avoid misleading and conflicting results. In the presence of random error, systematic error and lower power, potential limitations also include attenuated estimates, especially in the use of

Table 5

Estimated usual weighted mean and bootstrapped standard errors (B=500) for nutrients of public health concern (sodium, added sugar, saturated fat), per 1 unit increase in usual intakes in relation to all-cause mortality, 2004/2005 to 2011 (N=16,212 participants with 4,901 second recall; ≥ 20 years)

	Males		Females		Males			Females		
						95% Confidence interval			95% Confidence interval	
Variables	Mean	Standard error	Mean	Standard error	Hazard ratio per 1 unit	Lower	Upper	Hazard ratio per 1 unit	Lower	Upper
Sodium density, gr/1000 kcal (per 100 unit increase)	1,514.3	16.0	1,532.7	12.3	1.06	0.95	1.17	F	F	F
Added sugar,% (per 5% increase)	10.0	0.2	9.5	0.2	1.08	0.90	1.30	1.32	1.10	1.57
Saturated fat,% (per 5% increase)	10.2	0.1	10.2	0.1	1.50	0.97	2.30	1.04	0.94	1.14

F too unreliable to be published

Note: After winsorizing outliers, at baseline with adjustment for age, sequence of dietary recall and weekend or weekday effect, and age-adjusted hazard ratios and 95% confidence intervals.

Sources: Canadian Community Health Survey – Nutrition 2004; Canadian Vital Statistics – Death Database.

repeated dietary recalls. Although within-individual random errors do not always attenuate estimates.[51,52] Specifically, there are time-varying covariates for which the link with mortality may be weaker as time progresses away from the baseline observation in 2004.

Data resource access

The linked data are available through Statistics Canada's Research Data Centres.

Funding

Authors acknowledge funding from Statistics Canada, Health Canada and the Canadian Institutes of Health Research (CIHR) for the linkage project. This study was supported by a grant from Statistics Canada.

M.J. was supported by the CIHR Banting Postdoctoral Fellowships program, a CIHR Fellowship and Banting Foundation Award.

The record linkage was conducted by Statistics Canada with funding support from Health Canada, the Institute for Clinical Evaluative Sciences and the Ottawa Hospital Research Institute.

Acknowledgement

The authors would like to thank the linkage team with the Health Analysis Division, Statistics Canada.

Conflict of interest

The authors declare no conflict of interest.

References

1. Sanmartin C, Decady Y, Trudeau R, Dasylva A, Tjepkema M, Fines P, et al. Linking the Canadian Community Health Survey and the Canadian Mortality Database: An enhanced data source for the study of mortality. Health Reports. 2016; 27(12): 10-8.
2. Gray L, Batty GD, Craig P, Stewart C, Whyte B, Finlayson A, et al. Cohort profile: the Scottish health surveys cohort: linkage of study participants to routinely collected records for mortality, hospital discharge, cancer and offspring birth characteristics in three nationwide studies. International Journal of Epidemiology 2010; 39(2): 345-50.
3. Sanmartin C, Reicker A, Dasylva A, Rotermann M, Jeon SH, Fransoo R, et al. Data resource profile: The Canadian Hospitalization and Taxation Database (C-HAT). International Journal of Epidemiology. 2018; 47(3): 687.
4. Mindell J, Biddulph JP, Hirani V, Stamatakis E, Craig R, Nunn S, et al. Cohort profile: the health survey for England. International Journal of Epidemiology. 2012; 41(6): 1585-93.
5. Holman CD, Bass AJ, Rosman DL, Smith MB, Semmens JB, Glasson EJ, et al. A decade of data linkage in Western Australia: strategic design, applications and benefits of the WA data linkage system. Australian Health Review. 2008; 32(4): 766-77.
6. Winkler WE. Record linkage. In: Pfeffermann D, Rao CR, editors. Sample Surveys: Design, Methods and Applications. North Holland: Elsevier; 2009. p. 351-80.
7. Rotermann M, Sanmartin C, Carriere G, Trudeau R, St-Jean H, Saidi A, et al. Two approaches to linking census and hospital data. Health Reports. 2014; 25(10): 3-14.
8. Wilkins K, Shields M, Rotermann M. Smokers' use of acute care hospitals—a prospective study. Health Reports. 2009; 20(4): 75-83.
9. Bailly A CD, Fransoo R, Martens P, McDougall C, McGowan KL, Prior H, Schultz J, Soodeen RA. Adult Obesity in Manitoba: Prevalence, Associations, and Outcomes. Winnipeg (Canada): Manitoba Centre for Health Policy; 2011. Available from: https://www.hopkinsacg.org/document/adult-obesity-in-manitoba-prevalence-associations-and-outcomes-3rd-edition/.
10. Davidsen M, Kjoller M, Helweg-Larsen K. The Danish National Cohort Study (DANCOS). Scandinavian Journal of Public Health. 2011; 39(7 Suppl): 131-5.
11. Teckle P, Hannaford P, Sutton M. Is the health of people living in rural areas different from those in cities? Evidence from routine data linked with the Scottish Health Survey. BMC Health Services Research. 2012; 12: 43.
12. Hamer M, Batty GD, Stamatakis E, Kivimaki M. Comparison of risk factors for fatal stroke and ischemic heart disease: a prospective follow up of the health survey for England. Atherosclerosis. 2011; 219(2): 807-10.
13. Wilkins R, Tjepkema M, Mustard C, Choiniere R. The Canadian census mortality follow-up study, 1991 through 2001. Health Reports. 2008; 19(3): 25-43.
14. Hwang SW, Wilkins R, Tjepkema M, O'Campo PJ, Dunn JR. Mortality among residents of shelters, rooming houses, and hotels in Canada: 11 year follow-up study. BMJ. 2009; 339: b4036.
15. Tjepkema M, Wilkins R. Remaining life expectancy at age 25 and probability of survival to age 75, by socio-economic status and Aboriginal ancestry. Health Reports. 2011; 22(4): 31-6.
16. Ng E. The healthy immigrant effect and mortality rates. Health Reports. 2011; 22(4): 25-9.
17. Tjepkema M, Wilkins R, Senecal S, Guimond E, Penney C. Potential years of life lost at ages 25 to 74 among Metis and non-Status Indians, 1991 to 2001. Health Reports. 2011; 22(1): 37-46.
18. Dankovchik J, Hoopes MJ, Warren-Mears V, Knaster E. Disparities in life expectancy of pacific northwest American Indians and Alaska natives: analysis of linkage-corrected life tables. Public Health Reports. 2015; 130(1): 71-80.
19. Kinner SA, Degenhardt L, Coffey C, Hearps S, Spittal M, Sawyer SM, et al. Substance use and risk of death in young offenders: a prospective data linkage study. Drug Alcohol Rev. 2015; 34(1): 46-50.
20. Nambiar D, Weir A, Aspinall EJ, Stoove M, Hutchinson S, Dietze P, et al. Mortality and cause of death in a cohort of people who had ever injected drugs in Glasgow: 1982–2012. Drug Alcohol Dependence. 2015; 147: 215-21.
21. Manuel DG, Perez R, Bennett C, Rosella L, Taljaard M, Roberts M, et al. Seven More Years: The impact of smoking, alcohol, diet, physical activity and stress on health and life expectancy in Ontario. An ICES/PHO Report. Toronto: Institute for Clinical Evaluative Sciences and Public Health Ontario. 2012. Available from: https://www.publichealthontario.ca/-/media/documents/S/2012/seven-more-years.pdf?la=en.
22. Health Canada. Canadian Community Health Survey, Cycle 2.2, Nutrition (2004): A Guide to Accessing and Interpreting the Data. 2006. Available from: http://www.hc-sc.gc.ca/fn-an/surveill/nutrition/commun/cchs_guide_escc-eng.php
23. Health Canada. Reference Guide to Understanding and Using the Data. 2015 Canadian Community Health Survey – Nutrition. 2017. Available from: https://www.canada.ca/en/health-canada/services/food-nutrition/food-nutrition-surveillance/health-nutrition-surveys/canadian-community-health-survey-cchs/reference-guide-understanding-using-data-2015.html.
24. Statistics Canada. Canadian Community Health Survey – Nutrition (CCHS) 2015 2014. Available from: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5049
25. Moshfegh AJ BL, Perloff B, LaComb R,. Improved method for the 24-hour dietary recall for use in national surveys [abstract]. FASEB J. 1999; 13: A603.

26. United States Department of Agriculture Agricultural Research Service. USDA automated multiple-pass method [Internet]. 2009 [cited 2016 Oct 02] Available from: http://www.ars.usda.gov/Services/docs.htm?docid=7710.

27. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. The Journal of Nutrition. 2006; 106(10): 1575-87.

28. National Cancer Institute. Usual Dietary Intakes: The NCI Method. Cited 17 May 2018. Available from: https://epi.grants.cancer.gov/diet/usualintakes/method.html [Internet]. [cited.

29. Health Canada. The Canadian nutrient file. Ottawa (Canada): Nutrition Research Division, Food Directorate; 2001.

30. Statistics Canada. Canadian Vital Statistics – Death Database. Available from: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3233.

31. Fellegi IP, Sunter AB. A theory for record linkage. J Am Stat Assoc. 1969; 64(328): 1183-210.

32. Statistics Canada. G-Link Version 3.0 User Guide (internal document).

33. Statistics Canada. Approved Record Linkages.

34. Statistics Canada. Directive on Record Linkage. 2011. Available from: http://www.statcan.gc.ca/record-enregistrement/policy4-1-politique4-1-eng.htm.

35. Statistics Canada. Trends in Mortality Rates 2000–2011. 2015. Available from: http://www.statcan.gc.ca/pub/82-625-x/2014001/article/11897-eng.htm.

36. Statistics Canada. Deaths, by age and sex, Canada. CANSIM Table 102-0503. Available from: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310070901.

37. Statistics Canada. Deaths and mortality rates, by age group and sex, Canada, provinces and territories. CANSIM Table 102-0504. Available from: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310071001.

38. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. American Journal of Epidemiology. 1986; 124(1): 17-27.

39. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. American Journal of Clinical Nutrition. 1997; 65(4 Suppl): 1220S-8S; discussion 9S-31S.

40. Brisbois TD, Marsden SL, Anderson GH, Sievenpiper JL. Estimated intakes and sources of total and added sugars in the Canadian diet. Nutrients. 2014; 6(5): 1899-912.

41. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. Journal National Cancer Institute. 2011; 103(14): 1086-92.

42. Tooze JA, Kipnis V, Buckman DW, Carroll RJ, Freedman LS, Guenther PM, et al. A mixed-effects model approach for estimating the distribution of usual intake of nutrients: the NCI method. Statistics in Medicine. 2010; 29(27): 2857-68.

43. Jessri M, Ng AP, L'Abbe MR. Adapting the Healthy Eating Index 2010 for the Canadian Population: Evidence from the Canadian National Nutrition Survey. Nutrients. 2017; 9(8).

44. Jessri M, Wolfinger RD, Lou WY, L'Abbe MR. Identification of dietary patterns associated with obesity in a nationally representative survey of Canadian adults: application of a priori, hybrid, and simplified dietary pattern techniques. American Journal of Clinical Nutrition. 2017; 105(3): 669-84.

45. Jessri M, Nishi SK, L'Abbe MR. Assessing the nutritional quality of diets of Canadian children and adolescents using the 2014 Health Canada Surveillance Tool Tier System. BMC Public Health. 2016; 16: 381.

46. Jessri M, Nishi SK, L'Abbe MR. Assessing the Nutritional Quality of Diets of Canadian Adults Using the 2014 Health Canada Surveillance Tool Tier System. Nutrients. 2015; 7(12): 10447-68.

47. Jessri M, Lou WY, L'Abbe MR. The 2015 Dietary Guidelines for Americans is associated with a more nutrient-dense diet and a lower risk of obesity. American Journal of Clinical Nutrition. 2016; 104(5): 1378-92.

48. Mahsa Jessri, Deirdre Hennessey, Anan Bader Eddeen, Carol Bennett, Zefeng Zhang, Quanhe Yang, et al. Sodium, Added Sugar and Saturated Fat Intake in Relation to Mortality and Cardiovascular Disease Events in Adults: Canadian National Nutrition Survey Linked with Vital Statistics and Health Administrative Databases. British Journal of Nutrition. 2022; Accepted: 1-33.

49. Keogh RH, Shaw PA, Gustafson P, Carroll RJ, Deffner V, Dodd KW, et al. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 1–Basic theory and simple methods of adjustment. Statistics in Medicine. 2020; 39(16): 2197-231.

50. Jessri M, Lou YW, L'Abbe MR. Evaluation of different methods to handle misreporting in obesity research: Evidence from the Canadian national nutrition survey. Br J Nutr. 2016; 115(1): 147-59.

51. van Smeden M, Lash TL, Groenwold RHH. Reflection on modern methods: five myths about measurement error in epidemiological research. International Journal of Epidemiology. 2020; 49(1): 338-47.

52. Brakenhoff TB, van Smeden M, Visseren FLJ, Groenwold RHH. Random measurement error: Why worry? An example of cardiovascular risk factors. PLoS One. 2018; 13(2): e0192298.