Validation of Canproj for Projecting Canadian Cancer Incidence Rates

Public Health Agency of Canada / Agence de la santé publique du Canada

A Demers, Z Qiu, R Dewar, A Shaw

Centre for Surveillance and Applied Research, Public Health Agency of Canada

Introduction

Every year, the Public Health Agency of Canada works with the Canadian Cancer Society and Statistics Canada to publish a comprehensive report describing cancer incidence and mortality in Canada by age, sex, and geography. The report includes a series of count and rate projections that fill the gap between the latest available year of data and the year the report is released.

In an effort to further advance cancer predictions in Canada, Qiu et al. developed the Canproj package that includes three key innovations to Nordpred:

1. The replacement of the Poisson distribution by negative-binomial distribution when over-dispersion is present,
2. An age-cohort model, and
3. A set of hybrid models that combine the strengths of Poisson or negative-binomial regression, the joinpoint method, and an average method for projections based on age-specific counts.

Objective

The objective of this project was to validate the Canproj model for the short-term projection of cancer rates. Specifically, we compared the accuracy of Canproj models using cross validation techniques and graphical representation and evaluated the automatic model selection features of Canproj to assess the capacity of these functions to select a good model.

Methods

Incident cases were abstracted from the Canadian Cancer Registry (CCR) and National Cancer Incidence Reporting System (NCIRS) were used for the analysis. Projection models included Nordpred, the age-cohort model, hybrid models incorporating age and potentially period effects (age-specific or all ages), and the 5-year average model.

Methods

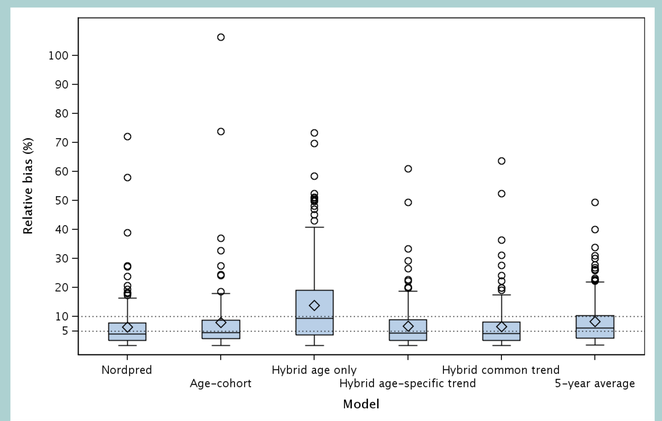
Cross-validation was used to estimate the expected level of fit. The data were split in two unequal subsamples; data from 2011 to 2014 was the independent testing data and the data from 1986 to 2010 was the training data.

We used the relative bias (RB) as measure of fit to compare the observed age-standardized rates (O) and the estimated ones (E);

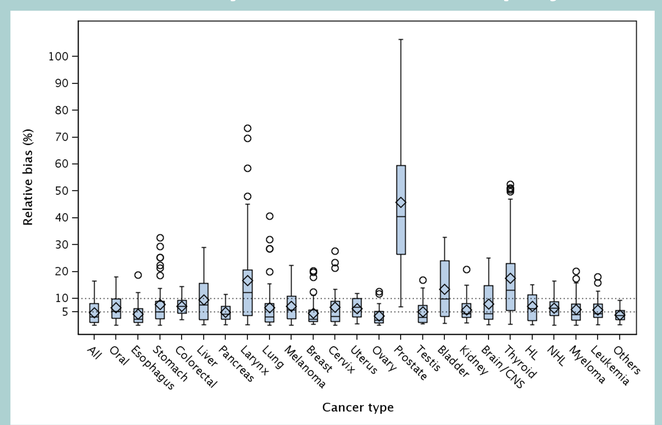
$$RB_t = \frac{|E_t - O_t|}{O_t} \times 100 \text{ , where } t = 2011 \text{ to } 2014$$

Results

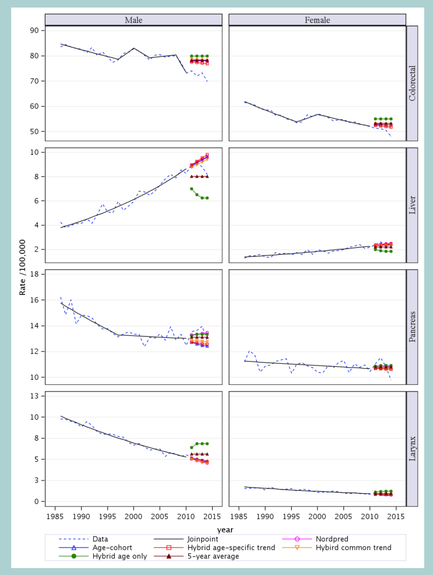
Relative bias by projection model for all cancers sites



Relative bias by cancer site for all projection models



Actual and projected age-standardized rates by sex and cancer site



Canproj decision tree: average relative bias by model, sex and cancer site

Sex	Cancer type	Nordpred	Age-cohort	Hybrid: age-specific	Hybrid: common trend	Hybrid: age only	5 year average
Male	All cancers	**10.9**	10.1	5.9	**5.7**	7.8	7.5
	Oral	**8.6**	13.2	14.8	12.6	**1.6**	5.9
	Esophagus	3.9	**2.6**	**4.2**	3.2	6.8	3.0
	Stomach	**3.4**	3.7	4.1	3.7	26.1	9.5
	Colorectal	**7.9**	8.2	**6.9**	8.6	10.7	8.3
	Liver	**6.8**	6.5	7.6	**6.2**	26.0	8.6
	Pancreas	3.2	7.0	**6.2**	5.0	**3.2**	3.8
	Larynx	**2.1**	3.6	2.5	2.2	39.4	16.2
	Lung and bronchus	**3.2**	1.7	2.6	**1.6**	32.4	12.9
	Melanoma	**2.1**	5.5	2.8	3.9	17.6	7.8
	Breast	**6.2**	6.8	8.6	**8.3**	6.4	6.2
	Prostate	**45.9**	86.4	38.9	41.8	31.3	**30.6**
	Testis	2.4	**2.4**	2.2	**2.1**	14.3	7.1
	Urinary bladder	**12.3**	17.3	13.7	15.9	**11.2**	11.6
	Kidney and renal pelvis	**5.5**	3.0	2.4	**2.3**	9.4	5.3
	Brain/CNS	**6.7**	**6.0**	7.0	6.3	10.0	8.0
	Thyroid	**3.3**	16.8	13.4	13.5	46.8	26.9
	Hodgkin lymphoma	**3.4**	3.4	3.5	3.4	**5.5**	3.8
	Non-Hodgkin lymphoma	**7.5**	7.2	**6.9**	7.0	8.5	7.5
	Myeloma	6.6	**6.3**	**6.6**	6.4	10.7	7.9
	Leukemia	6.0	3.9	**6.0**	5.5	**2.8**	3.6
	All others	**3.8**	4.8	3.8	3.6	4.6	**3.6**
Female	All cancers	**1.3**	1.1	**1.0**	1.1	3.4	1.1
	Oral	3.4	**4.1**	4.4	3.9	**2.9**	3.3
	Esophagus	3.6	3.5	3.2	**3.0**	8.8	3.9
	Stomach	**2.7**	2.9	7.2	4.1	20.7	4.9
	Colorectal	**5.2**	4.9	**3.7**	5.4	9.5	6.1
	Liver	**5.4**	**4.9**	5.2	5.5	21.8	9.0
	Pancreas	4.9	5.1	5.2	5.0	**4.9**	**4.7**
	Larynx	**10.2**	14.5	16.7	15.3	53.4	24.0
	Lung and bronchus	**1.6**	1.6	10.3	5.8	3.5	2.4
	Melanoma	**3.0**	8.4	4.2	4.1	16.9	9.1
	Breast	**2.3**	2.9	**1.2**	1.5	1.8	1.7
	Cervix uteri	**3.4**	4.6	1.5	**1.5**	20.9	8.3
	uterus	**2.8**	7.9	3.0	3.6	10.1	10.1
	Ovary	**1.9**	**1.8**	1.8	1.8	9.7	4.0
	Urinary bladder	**14.6**	14.7	13.7	15.0	**10.7**	10.7
	Kidney and renal pelvis	**12.0**	5.4	6.0	5.8	7.7	**4.1**
	Brain/CNS	**7.8**	**7.1**	8.2	7.7	10.8	9.1
	Thyroid	**4.3**	5.7	5.5	5.3	47.9	21.6
	Hodgkin lymphoma	10.1	**10.5**	11.5	11.0	**7.9**	10.0
	Non-Hodgkin lymphoma	**5.6**	5.3	5.4	**5.3**	5.8	5.7
	Myeloma	**3.7**	3.9	**3.8**	3.8	6.4	4.4
	Leukemia	**13.2**	5.6	5.2	6.4	**4.1**	5.1
	All others	**3.0**	4.2	3.7	4.0	3.1	3.2

Light green cells are the projection models selected by Canproj; lilac cells are the models with smallest relative bias; dark green cells indicate that Canproj selection is the model with the smallest relative bias

2010-2014 median relative bias (%) by model and cancer type

Sex	Cancer type	Nordpred	Age-cohort	Hybrid: Age-specific trend	Hybrid: Common trend	Hybrid: Age only	5-year average	Diagnostic: JP	Diagnostic: RB ratio	Diagnostic: RB (%)
Male	All cancers	11.0	10.1	5.9	**5.6**	7.8	7.5	**2007**	**5.7**	1.0
	Oral	8.8	13.4	14.8	12.7	**1.0**	6.3	**2003**	0.7	1.3
	Esophagus	2.6	**2.2**	3.0	2.7	5.9	2.2	**2005**	0.8	2.8
	Stomach	**3.1**	3.8	3.9	3.9	25.2	8.5	1986	1.6	1.9
	Colorectal	7.4	7.7	**6.4**	8.1	10.2	7.8	**2008**	**16.0**	0.4
	Liver	4.4	4.2	4.8	**3.9**	26.2	10.0	1986	1.0	4.0
	Pancreas	**3.0**	6.9	6.5	5.4	3.1	3.5	1997	**2.3**	1.3
	Larynx	**1.4**	3.6	2.4	2.0	44.0	18.2	1986	0.6	2.4
	Lung and bronchus	3.1	1.8	1.9	**1.7**	30.2	11.1	1986	1.3	1.3
	Melanoma	**1.4**	6.1	1.5	4.6	17.6	8.3	1986	0.7	2.0
	Breast	4.9	**4.3**	6.3	6.0	6.6	6.7	1986	0.6	6.9
	Prostate	48.4	90.1	41.4	44.4	33.8	**33.1**	**2001**	**11.2**	2.9
	Testis	1.4	1.4	1.3	**1.3**	13.9	7.2	1986	0.3	4.4
	Urinary bladder	10.1	15.9	12.4	14.7	9.8	**9.3**	1990	**3.5**	2.7
	Kidney and renal pelvis	5.0	2.8	2.0	**2.0**	9.1	4.8	1998	1.0	1.9
	Brain/CNS	4.1	**2.9**	4.3	3.4	7.7	5.6	1986	1.5	2.0
	Thyroid	**3.7**	17.0	13.0	13.0	48.9	27.4	1997	0.8	4.6
	Hodgkin lymphoma	1.4	1.4	2.1	1.6	3.4	**1.3**	1986	0.5	2.6
	Non-Hodgkin lymphoma	7.7	7.4	**6.6**	7.3	8.5	7.8	**2007**	**4.4**	1.5
	Myeloma	5.2	4.7	5.1	**4.6**	10.0	6.8	1986	1.3	3.6
	Leukemia	6.2	3.8	6.1	5.2	**0.8**	3.5	1994	0.4	2.0
	All others	4.0	4.1	2.8	**2.8**	4.5	3.6	**2003**	**2.1**	1.3
Female	All cancers	0.9	**0.8**	0.8	0.8	3.3	0.9	1986	1.0	0.8
	Oral	3.1	4.2	4.2	4.2	**1.7**	2.8	1986	0.6	2.9
	Esophagus	1.1	1.1	1.5	**0.9**	6.8	1.3	1986	0.3	3.4
	Stomach	**1.3**	2.7	7.2	4.4	20.5	3.5	1992	0.6	2.0
	Colorectal	4.0	3.7	**2.7**	4.2	8.3	4.9	2000	**5.5**	0.5
	Liver	5.3	4.8	5.1	**4.4**	21.2	8.9	1986	0.7	6.4
	Pancreas	4.3	5.1	5.5	4.9	**3.5**	4.2	1986	1.4	2.5
	Larynx	**10.0**	13.6	17.3	15.1	64.1	28.5	1986	1.8	5.5
	Lung and bronchus	1.3	**1.3**	9.0	4.9	3.9	1.9	**2006**	**2.1**	0.6
	Melanoma	**2.8**	8.4	3.5	3.4	16.3	9.3	1992	1.4	2.0
	Breast	2.3	3.2	**0.7**	1.3	1.9	1.6	1991	0.4	1.9
	Cervix uteri	3.4	4.7	**1.4**	1.4	22.3	8.6	**2006**	0.7	2.1
	uterus	3.1	8.0	3.4	**2.6**	10.0	10.0	**2005**	1.8	1.5
	Ovary	1.1	**1.1**	1.2	1.7	9.9	4.7	1986	0.6	1.7
	Urinary bladder	14.1	14.2	13.1	14.5	10.1	**9.7**	1986	**3.1**	3.1
	Kidney and renal pelvis	12.0	4.3	4.6	4.3	6.1	**4.2**	1986	1.8	2.4
	Brain/CNS	5.6	**5.3**	5.7	5.5	8.8	6.2	1986	**2.2**	2.4
	Thyroid	**4.5**	6.3	5.4	5.9	50.1	21.7	**2005**	**2.7**	1.7
	Hodgkin lymphoma	10.3	10.9	12.2	11.6	**8.0**	10.4	1986	**2.4**	3.4
	Non-Hodgkin lymphoma	5.3	4.9	**4.4**	4.9	5.8	6.0	1997	**3.9**	1.1
	Myeloma	**3.8**	4.1	3.9	3.9	6.8	4.2	1986	1.0	3.7
	Leukemia	14.3	5.0	4.7	6.8	**2.6**	4.8	**2001**	1.5	1.7
	All others	3.1	4.2	3.6	4.1	**3.0**	3.1	**2004**	**3.5**	0.8

Conclusions

Cancer projection methods have evolved over the years. The Canproj package, which includes key innovations not included in Nordpred, can provide accurate projections for most cancer sites in Canada, which can support future planning for cancer control strategies treatments and interventions.

Note: Quebec has not submitted cancer incidence data since 2010, and was not included in the current analysis.