Appendix 1 (as supplied by the authors):

Imputation of 2011-2015 Cancer Incidence Data for Quebec

Appendix to: Brenner DR, Ruan Y, Shaw E, et al. Age-standardized cancer-incidence trends in Canada, 1971–2015. *CMAJ* 2019. doi: 10.1503/cmaj.190355.

The cancer incidence for the province of Quebec was only available up to year 2010. In order to estimate the national trend of cancer incidence, we imputed the incidence data of Quebec from year 2011 to 2015, by sex and cancer type. We proposed 3 approaches to impute the cancer incidence data of Quebec in 2011 to 2015.

The first approach simply uses the average age-specific rates in the recent five years (i.e., 2006-2010) in Quebec, and assumes the constant rates in future incidence projections:

$$IR_{QC,age}^{yr\_f} = \frac{1}{5}\sum_{yr=2006}^{2010} IR_{QC,age}^{yr}$$

Where IR_{QC} is the incidence rate of a specific cancer in Quebec, *yr_f* is the projected years from 2011 to 2015, age indicates a specific 5-year age group from 20 and up. We named this approach "5-yr average".

The second approach employs Canproj, an R package that uses a decision tree algorithm to find the best fitting age-period-cohort model, to impute the cancer incidence of Quebec from 2011 to 2015, using the past cancer incidence data in Quebec from 1971 to 2010. The details of this method can be found in our previous publication [1]. We named this approach "APC model".

The third approach accounts for the available incidence data up to 2015 from Canada (excluding Quebec). A recent incidence trend is estimated from the national incidence data excluding Quebec, using a Poisson regression model. We arbitrarily used the data of the recent 8 years (2008 – 2015):

$$\log\left(IR_{sex,age,cancer}\right) = \beta_0 + \beta_1 Year$$

Where $IR_{sex,age}$ is the incidence rate in a specific sex-age-cancer stratum. β_1 is the incremental incidence rate ratio (IRR) by year. The estimated IRR was then applied to the incidence rate of Quebec in 2010 to impute the cancer incidence from 2011 to 2015. The assumption of this approach is that the recent incidence trend in Quebec is the same as that of Canada excluding Quebec. We named this approach "IRR fitting".

To evaluate the performance of the aforementioned approaches, we tested these methods on the historical cancer incidence data from 2001 to 2010. For instance, we used the subset of incidence data up to 1996 to project the incidence of 2001 in Quebec. The predicted incidence was then compared to the observed 2001 incidence data to evaluate the accuracy.

The performance was evaluated on age-aggregated incidence using four scores:

1) The fraction of predictions that are within 10% of the observed incidence. Higher score means more accurate imputations.

2) The fraction of predictions exceeding 25% of the observed incidence. Lower score means more reliable imputations.

3) The average absolute relative difference: $\frac{1}{n}\sum_{i=1}^{n}\frac{|I_{pred,\ i}-I_{obs,\ i}|}{I_{obs,\ i}}$, which measures the overall deviation in the imputations. n = 10 (year 2001 – 2010) for sex-specific cancers; n = 20 (year 2001 – 2010; male and female) for cancers in both sexes.

4) Overall average relative difference: $\frac{1}{n}\sum_{i=1}^{n}\frac{I_{pred,\ i}-I_{obs,\ i}}{I_{obs,\ i}}$, which measures the bias in the imputation method.

The scores by each cancer type were shown in Appendix Table 1. For each cancer type, we selected the imputation approach that gave high score 1, and gave score 2 – 4 that were most close to 0. For 14 out of 24 cancer types, the “IRR fitting” approach was selected as the best imputation method. In contrast, the “APC model” only outperformed the other two approaches in cervical and stomach cancer, and performed slightly better in oral and gallbladder cancer.

Appendix 1 Table 1: summary of the performance of the three imputation approaches by cancer types. The highlighted approaches were the approaches selected for imputing the cancer incidence of the specific cancer type.

Cancer type	Imputation approach	Score 1	Score 2	Score 3	Score 4
Bladder	**5-yr average**	17/20	0/20	4.50%	0.20%
	APC model	12/20	1/20	9.30%	-2.10%
	IRR fitting	13/20	0/20	8.30%	-7.50%
Breast	5-yr average	10/10	0/10	5%	3.50%
	APC model	5/10	0/10	9%	9%
	IRR fitting	10/10	0/10	3%	-0.40%
Cervix	5-yr average	4/10	1/10	12.20%	12.20%
	APC model	8/10	0/10	6.10%	-1.50%
	IRR fitting	8/10	0/10	7%	-4.60%
Colon	5-yr average	17/20	0/20	6.60%	6.20%
	APC model	16/20	0/20	6.40%	4.30%
	IRR fitting	19/20	0/20	3.10%	-1.10%
Colorectal	5-yr average	16/20	0/20	6.80%	5.90%
	APC model	11/20	0/20	9.20%	7.40%
	IRR fitting	20/20	0/20	2.80%	0.60%
Uterus	5-yr average	9/10	0/10	4.90%	1.70%
	APC model	8/10	0/10	5.30%	-4.10%
	IRR fitting	10/10	0/10	2.40%	1.30%
Gallbladder	5-yr average	3/20	13/20	35.70%	33.20%
	APC model	3/20	12/20	30.20%	27.60%
	IRR fitting	3/20	12/20	43.70%	31.20%
Kidney	5-yr average	16/20	0/20	7.20%	-7.10%
	APC model	12/20	1/20	9.20%	-1.80%
	IRR fitting	16/20	0/20	7.30%	1.60%
Larynx	5-yr average	1/20	14/20	36.70%	36.20%
	APC model	8/20	7/20	28.50%	-9.80%
	IRR fitting	8/20	5/20	18.90%	11%
Liver	5-yr average	5/20	0/20	14.70%	-13.70%
	APC model	6/20	2/20	21.90%	-8.10%
	IRR fitting	8/20	1/20	12.10%	6.30%
Lung	5-yr average	1/20	1/20	15.70%	3.30%
	APC model	0/20	4/20	20.90%	20.90%
	IRR fitting	17/20	0/20	5.40%	-2%
Melanoma	5-yr average	8/20	4/20	14.80%	-12.60%
	APC model	7/20	4/20	16.30%	7.50%
	IRR fitting	7/20	1/20	12.20%	-0.80%
Myeloid leukemia	5-yr average	9/20	3/20	15%	11.30%
	APC model	5/20	8/20	25.60%	19.30%
	IRR fitting	7/20	4/20	17%	5.10%
Myeloma	**5-yr average**	15/20	0/20	8.30%	3.80%
	APC model	8/20	3/20	14%	11.40%
	IRR fitting	10/20	2/20	11.80%	5%
Non-Hodgkin lymphoma	**5-yr average**	19/20	0/20	4%	2.80%
	APC model	14/20	1/20	12.30%	-1.90%
	IRR fitting	16/20	0/20	6.20%	4.70%
Esophagus	**5-yr average**	12/20	2/20	10.70%	7%
	APC model	8/20	6/20	19.40%	18.20%
	IRR fitting	9/20	4/20	14.70%	4.70%
Oral	5-yr average	15/20	1/20	8.70%	6.50%
	APC model	14/20	0/20	7.90%	-1.90%
	IRR fitting	9/20	2/20	10.80%	1.10%
Ovary	5-yr average	8/10	0/10	8.40%	8.40%
	APC model	7/10	0/10	8.70%	8.60%

	IRR fitting	9/10	0/10	4.40%	0.20%
Pancreas	5-yr average	15/20	1/20	8.40%	7.40%
	APC model	2/20	6/20	22.60%	22.60%
	IRR fitting	15/20	1/20	8.10%	3.80%
Prostate	**5-yr average**	8/10	0/10	7.40%	0.50%
	APC model	2/10	4/10	25%	19.30%
	IRR fitting	8/10	0/10	6.30%	-0.90%
Rectum	5-yr average	15/20	0/20	7.10%	5.70%
	APC model	11/20	1/20	13.10%	-0.30%
	IRR fitting	18/20	0/20	5.50%	4.40%
Small Intestine	**5-yr average**	6/20	4/20	16.40%	-9.30%
	APC model	4/20	8/20	32.10%	28%
	IRR fitting	2/20	8/20	30%	21.20%
Stomach	5-yr average	0/20	15/20	31.20%	31.20%
	APC model	13/20	3/20	11%	1.60%
	IRR fitting	13/20	2/20	9.20%	7.10%
Thyroid	5-yr average	0/20	20/20	43%	-43%
	APC model	4/20	5/20	19.20%	-17.70%
	IRR fitting	14/20	1/20	8.70%	-4%

Reference:

[1] Poirier, AE, Ruan Y, Walter SD, Franco EL, Villeneuve PJ, King WD, Volesky KD, O'Sullivan DE, Friedenreich CM, Brenner DR; ComPARe Study Team. The future burden of cancer in Canada: Long-term cancer incidence projections 2013-2042. *Cancer Epidemiology*, 2019, 59:199-207.