Received: 28 August 2018 | Revised: 7 January 2019 | Accepted: 16 January 2019

DOI: 10.1002/hed.25682

WILEY

ORIGINAL ARTICLE

Developing case-finding algorithms for second events of oropharyngeal cancer using administrative data: A population-based validation study

Yuan Xu MD, PhD[1,2,3] | Shiying Kong MSc[1,2,3] | Winson Y. Cheung MD, MPH[2,3] | May Lynn Quan MD, MSc[1,2,3] | Steven C. Nakoneshny BSc[4] | Joseph C. Dort MD, MSc[1,2,3,4]

[1]Department of Surgery, Foothills Medical Centre, University of Calgary, Calgary, Alberta, Canada

[2]Department of Community Health Science, Foothills Medical Centre, University of Calgary, Calgary, Alberta, Canada

[3]Department of Oncology, University of Calgary, Tom Baker Cancer Centre, Calgary, Alberta, Canada

[4]The Ohlson Research Initiative, Arnie Charbonneau Cancer Institute, University of Calgary, Calgary, Alberta, Canada

Correspondence
Yuan Xu, Department of Surgery, Foothills Medical Centre, University of Calgary, 3280 Hospital Drive NW, Calgary, Alberta, Canada T2N4Z6.
Email: yuxu@ucalgary.ca

Abstract

Background: Second event (recurrence or second primary cancer)-free survival is an important indicator for assessing treatment efficacy. However, second events are not explicitly documented in administrative data such as cancer registries. Thus, validated algorithms using administrative data are needed to identify second events of oropharyngeal cancers.

Methods: The algorithms were developed using classification and regression tree models. Data from chart review served as the reference standard. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were calculated.

Results: The high-sensitivity algorithm achieved 87.9% (95% confidence interval: 82.2%-93.6%) sensitivity, 84.5% (81.1%-87.8%) specificity, 61.2% (54.1%-68.4%) PPV, 96.2% (94.2%-98.1%) NPV, and 85.2% (82.3%-88.1%) accuracy. The high-PPV algorithm obtained 52.4% (43.6%-61.2%) sensitivity, 99.1% (98.2%-100.0%) specificity, 94.2% (88.7%-99.7%) PPV, 88.2% (85.3%-91.0%) NPV, and 88.9% (86.3%-91.5%) accuracy.

Conclusion: The validity of the algorithms for identifying second events following primary treatment of oropharyngeal cancers was acceptable.

KEYWORDS

administrative data, case-finding algorithm, oropharyngeal cancer recurrence, population-based study, validation study

1 | INTRODUCTION

Although there have been significant treatment advances for oropharyngeal cancers in recent years, recurrence rates remain high. Therefore, cancer-free survival continues to be a clinically important indicator in outcomes studies.[1,2] However, second events from cancer, such as recurrence or new primary, are not typically recorded in routinely collected population-based administrative data sources, such as the cancer registry.[3,4] Thus, ascertainment of second events must usually rely on chart review, which is expensive and inefficient. This can be a barrier that prevents researchers from conducting large-scale studies of cancer-free survival, which are necessary for comparing treatment efficacy, assessing health care quality, and informing decision making for patients with cancer.

In Canada, where a universal health care system exists, there are comprehensive administrative data at both the provincial and national levels that cover almost all medical encounters throughout a patient's disease trajectory. Using these administrative data, the pattern of medical care that a patient receives may actually indicate second events, because recurrences or a new primary often require intensive health care resources, such as reoperations, additional

chemotherapy and/or radiation, or other health care services after the initial cancer treatment. Understanding and characterizing these unique patterns of health care use represents a potential strategy for identifying second events in oropharyngeal cancers using administrative data.

One study[5] explored the use of treatment data from one cancer center in the United Kingdom to identify recurrences of head and neck cancers and reported that their algorithm had moderate validity; however, the sample size of this previous study was small and not population based. More importantly, the algorithm was not applicable to Canada because of the differences in data availability and data systems. The present study aims to develop algorithms for identifying second events in oropharyngeal cancer using routinely collected administrative data that are available in Canada and other similar jurisdictions.

2 | METHODS

2.1 | Data and study cohort

This is a population-based, retrospective validation study using provincial administrative data from Alberta, Canada. The study included all adult patients who were diagnosed with oropharyngeal cancer between January 1, 2009 and December 31, 2015 and who were treated at the Tom Baker Cancer Center, which is the only center providing treatment for patients with oropharyngeal cancer in southern Alberta including Calgary and surrounding regions. The end of observation date was up to September 31, 2017 or the last follow-up date (died, move out of the area, or diagnosis of a second primary nonoropharyngeal tumor). Except for the patients who were diagnosed with a second primary nonoropharyngeal tumor, censored patients were included in the analyses given that majority (96%) of the patients had a follow-up more than 6 months. Patients with distant metastasis, multiple primary tumors, or who did not undergo curative intent treatment were excluded. Figure 1 displays the flowchart for cohort selection. This study was approved by the Health Research Ethics Board of Alberta—Cancer Committee (reference number: HREBA.CC-16-0644). As this study only used the secondary data of patients, the consent to participate was not required by the ethics review committee. For development of the algorithms, we retrieved administrative data from the physician claims, hospital discharge abstract, ambulatory care, and the cancer registry and vital statistics data up to September 31, 2017. These datasets were linked by a unique patient health number. Data from a prospectively collected chart review served as the gold standard for validating the developed algorithms. Chart review and abstraction were performed by a single investigator (SN). SN underwent extensive training under the direct supervision of a senior surgeon (JD) after which time SN worked independently. Cases in which data were unclear or

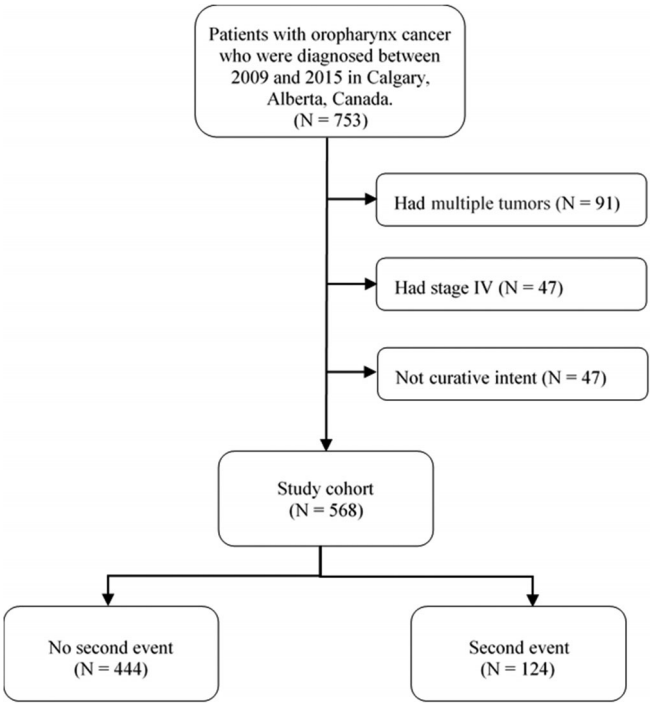

FIGURE 1 The flowchart of selection of study cohort

ambiguous at the time of primary data entry were reviewed in concert by SN and JD with outcomes agreed upon by consensus. All data were cleaned and validated by SN prior to analysis by YX and SK. Ambiguous data during analysis were resolved by consensus of the investigative team.

Because Alberta has a publicly funded, single-payer health care system, all patient medical encounters (such as hospitalizations, emergency room and clinic visits, and surgeries) and endpoint status (ie, alive or deceased) following initial cancer treatment are captured in various administrative data sources, including the Alberta Cancer Registry, Discharge Abstract Data, National Ambulatory Care Reporting System, physician billing claims, and vital statistics. Information on chemotherapy and radiation are available from the Alberta Cancer Measurement Outcomes Research and Evaluation unit, which integrates all available data related to cancer treatments across the provinces.

Second events were ascertained by primary chart review and served as the gold standard to validate the developed algorithms. A second event was defined as a recurrence or second primary cancer at local, regional, or distant sites from the index primary tumor at least 180 days after initial curative intent treatment.

2.2 | Definition of study variables

We assumed that patients with oropharyngeal cancer who received subsequent chemotherapy or radiation, or who underwent a diagnostic or surgical procedure, or who had frequent (or a change in the frequency of) visits to the cancer center beyond the initial treatment period would have a higher probability of a cancer recurrence than those who did

TABLE 1 study variables specifications and ICD codes

Procedure and diagnosis	Codes
Surgery/procedure	
Neck dissection	CCI: 1.GE.91.^
	Alberta physician billing codes: 52.31^
Partial/hemiglossectomy	CCI: 1.FJ.87.^
	Alberta physician billing codes: 37.1A, 37.1B, 37.2
Orbital exenteration	CCI: 1.CP.91^
	Alberta physician billing codes: 29.4A
Lip excision	CCI: 1.YE.87.^
	Alberta physician billing codes: 98.6A, 98.6B
Tonsillectomy	CCI: 1.FR.89.^
	Alberta physician billing codes: 43.1A
Laryngectomy	CCI: 1.GE.87.^
	Alberta physician billing codes: 42.3A, 42.3B, 42.3C
Maxillectomy	CCI: 1.ED.87^, 1.ED.91^
	Physician billing codes: 88.4A
Temporal bone subtotal resection	CCI: 1.DR.91^ Physician billing codes: 89.78E
Tracheotomy	CCI: 1.GJ.77^
	Physician billing codes: 43.1A
Parotidectomy	CCI: 1.FM.87^, 1.FM.89^, 1.FM.91^
	Physician billing codes: 38.23A
Submandibular gland resection	CCI: 1.FN.87^, 1.FN.89^
	Physician billing codes: 38.21A
Laryngoscopy	CCI: 2.GE.70^, 2.GE.71^
	Physician billing codes: 42.09B, 42.09D
Primary cancer site (oropharyngeal cancers)	ICD-O: C00.^- C10.^and C14.^
	ICD-9:140.^- 149.^
	ICD-10: C00.^- C10.^and C14.^
Specialty visits (physician specialty visit only)	Based on the provider classification, we identified the type of the physician specialty including the oncologists and head and neck surgeons.
Death caused by cancer	ICD-9:140.^- 208.^
	ICD-10: C00.^- C97.^

Abbreviations: CCI, Canadian Classification of Health Intervention; ICD-9, International Classification of Disease—ninth edition; ICD-10, International Classification of Disease—tenth edition; ICD-O, International Classification of Disease for Oncology.
Note: Physician billing codes are derived from the Alberta Schedule of Medical Benefits (SOMB).

not. Therefore, we defined key input variables (the specifications and International Classification of Disease codes for the variables are presented in Table 1) for the development of the algorithms, as follows.

2.2.1 | A new round of chemotherapy

We counted the number of episodes of chemotherapy that occurred within a specific time frame after the primary treatment. An episode is defined as a separate administration of chemotherapy. We assessed various time frames after the primary treatment, including 180, 365, and 540 days, because patients with oropharyngeal cancer usually do not receive additional rounds of chemotherapy within 6 months following primary treatment. If the number of chemotherapy episodes was more than two, then we assumed that the patient experienced a second event of oropharyngeal cancer for the purposes of our algorithm.

2.2.2 | Occurrence of specific procedure and second treatments

We created indicator variables for the receipt of a procedure (eg, laryngoscopy) or a second treatment (eg, surgery or radiation therapy) following the primary cancer treatment. We created the indicator variable if such procedures or treatments occurred after 60, 90, 180, 365, or 540 days of the primary treatment.

2.2.3 | A new cluster of cancer center visits

We calculated the number of days in between two cancer center visits and determined if there was a new cluster of visits with an interval that exceeded a prespecified value (90, 120, or 180 days). Subsequently, we counted the number of visits within this cluster and created an indicator variable for whether the counts exceeded a prespecified value (such as three, four, or five encounters).

2.2.4 | Tumor features, patient characteristics, and other clinical factors

Considering that patients with oropharyngeal cancer with certain characteristics are more likely to recur, such as advanced tumor stage, we included patient and tumor characteristics and other potential indicators, which might be associated with a second event in the model, including patient's age at diagnosis, tumor stage, and death caused by cancer. All those variables were assessed in the algorithms but only the ones that contribute to the classification ability of the algorithms were included in the final algorithms.

2.3 | Statistical analyses

The study cohort was divided into a training (60%) and testing (40%) set. Development of the algorithms was conducted using classification and regression tree (CART) models that used the Monte Carlo simulation[6] to evaluate multiple parameters by accounting for all possible dichotomous cut-offs and interactions between the inputted variables. Moreover, CART permitted us to set the costs for misclassifications (such as false positives and false negatives), so we were able to develop different algorithms that optimized sensitivity, specificity, positive predictive value (PPV), or negative predictive value (NPV). We imported all created variables for algorithm development. The tree was split into two child nodes from the root (node) based on the input variables. All possible splits were evaluated and determined by minimizing the impurities among the input variables. Each interior node represented one variable. The splitting process was repeated until the stopping criteria were satisfied, that is, the number of observations in

the last node was less than 10, or there were no other splits available for the observations (ie, all observations in the subgroup were classified as the same outcome). In order to avoid overfitting the data and to simplify the algorithm, we selected the smallest tree with a misclassification rate within 1 SE of the rate of the optimal tree in which the misclassification rate was smallest among all subtrees.[7]

We evaluated the performance of the algorithm by comparing the predicted outcome with the reference standard chart review. Various validity metrics were calculated including sensitivity, specificity, PPV, NPV, and their corresponding 95% confidence intervals (CIs) based on an exact binomial distribution. Accuracy was also calculated to measure the overall misclassification rate which is equal to (true positive + true negative)/(true positive + false positive + true negative + false negative).

Considering that the algorithms may be used for different purposes[8–12] such as identifying a cohort of patients with second events to test a new treatment (ie, high PPV is desirable) or developing a surveillance program to follow second events (ie, high sensitivity is desirable), we developed a set of algorithms which prioritized sensitivity, PPV, or accuracy, respectively. To achieve this, we assessed various algorithms by selecting the input variables and adjusting the cost for misclassification to optimize the target metric (such as PPV) while maintaining a reasonable threshold for the other metrics (ie, sensitivity, specificity, and NPV).

To optimize the validity metrics, we tried to combine two algorithms together (ie, the high-sensitivity and high-PPV algorithms). This can result in discordant cases, so chart reviews were conducted to resolve the discrepancies.

All analyses were performed using CART (Salford Systems, San Diego, California) and SAS 9.4 (SAS Institute Inc., Cary, North Carolina).

3 | RESULTS

In total, 568 patients with oropharyngeal cancer who underwent curative intent treatment were included. After a median follow-up of 34 (range: 21-50) months, 124 (21.8%) developed second events (109 recurrence and 15 second primary oropharyngeal cancer) as ascertained by chart review. Of the 109 with recurrence, 83 (76.1%) were distant recurrences and 26 (23.8%) were local recurrences. Among the 15 with second primary patients with oropharyngeal cancer, 10 (66.7%) patients were stage IV, and the common locations were four (26.7%) tonsil, four (26.7%) tongue base, and seven (46.6%) overlapping sites. There were differences in the primary surgery, tumor stage, and mortality (overall and cancer specific) between patients with and without second events. Small differences were observed between two groups in terms of age ($P = 0.0509$) and chemotherapy ($P = 0.0724$). Otherwise, there were no significant differences in terms of sex, overall cancer stage, tumor grade, and primary radiation therapy between groups. Additional details are summarized in Table 2. In the final developed algorithms, the indicator variables for identifying second event included: age, a second surgery, a new round of chemotherapy or radiation therapy, death caused by cancer, and a second referral to oncologist 1 year after primary diagnosis (Figures 2–4).

As the validation results were similar in the testing set and in the entire cohort, we present the performance results of the algorithms in the entire cohort in Table 3 and Figures 2–4. The validity results for the testing set are displayed in the Supporting Information Table S1. The high-sensitivity algorithm achieved 87.9% (95% CI: 82.2%-93.6%) sensitivity, 84.5% (95% CI: 81.1%-87.8%) specificity, 61.2% (95% CI: 54.1%-68.4%) PPV, 96.2% (95% CI: 94.2%-98.1%) NPV, and 85.2% (95% CI: 82.3%-88.1%) accuracy. The high-PPV algorithm obtained 52.4% (95% CI: 43.6%-61.2%) sensitivity, 99.1% (95% CI: 98.2%-100.0%) specificity, 94.2% (95% CI: 88.7%-99.7%) PPV, 88.2% (95% CI: 85.3%-91.0%) NPV, and 88.9% (95% CI: 86.3%-91.5%) accuracy. The high-accuracy algorithm reached 73.4% (95% CI: 65.6%-81.2%) sensitivity, 96.2% (95% CI: 94.4%-98.0%) specificity, 84.3% (95% CI: 77.4%-91.1%) PPV, 92.8% (95% CI: 90.5%-95.2%) NPV, and 91.2% (95% CI: 88.9%-93.5%) accuracy.

The classification tree (Figure 2) for the algorithm with high sensitivity was interpreted as follows: the entire cohort (N = 568) was divided into two groups using the indicator variable "a new round of radiation therapy occurring 180 days after the primary treatment." In total, 45 patients who met this criterion were classified as a recurrent case. Patients (N = 523) who did not meet the above criterion were subsequently regarded as having a recurrence if they underwent a new round of chemotherapy at 180 days or more after the primary treatment. Using this criterion, 19 patients met our definition of a second event. For the remaining 504 patients, 27 were classified as second-event patients due to the receipt of another surgery 1 year after the primary treatment. This left 477 patients of whom 14 patients were referred to oncology 1 year or more after primary treatment; thus, they were defined as experiencing a second event. Finally, among the 463 patients who did not meet any of the above criteria, 73 patients were classified as second-event patients, because they were recorded as having died because of cancer.

The method of combining algorithms into one in order to achieve both high sensitivity and high PPV resulted in 19.2% (109 out of 568) discordant cases. If additional chart review was conducted to resolve these discrepant cases, the metrics would greatly improve to 87.9% (95% CI: 82.2%-93.6%) sensitivity, 99.1% (95% CI: 98.2%-100.0%) specificity, 96.5% (95% CI: 93.1%-99.9%) PPV, 96.7% (95% CI: 95.1%-98.3%), NPV and 96.7% (95% CI: 95.2%-98.1%) accuracy (Table 3).

We also validated the algorithms for identifying recurrence alone (rather than recurrence or second primary) and found that the performance of the algorithms was slightly lower than the performance of the algorithms for identifying

TABLE 2 The patient characteristics

	All patients (N = 568)	Patients without second event (N = 444)	Patients with second event (N = 124)	*P* value
Age, Mean (SD), y	59.4 (11.5)	58.9 (11.4)	61 (11.6)	0.07
Year of diagnosis				0.17
2009-2011	231 (40.7%)	174 (39.2%)	57 (46%)	
2012-2015	337 (59.3%)	270 (60.8%)	67 (54%)	
Sex				0.37
Female	146 (25.7%)	118 (26.6%)	28 (22.6%)	
Male	422 (74.3%)	326 (73.4%)	96 (77.4%)	
T classification[a]				0.005
T1	144 (25.4%)	126 (28.4%)	18 (14.5%)	
T2	214 (37.7%)	170 (38.3%)	44 (35.5%)	
T3	43 (7.6%)	31 (7%)	12 (9.7%)	
T4	157 (27.6%)	110 (24.8%)	47 (37.9%)	
Unknown	10 (1.8%)	7 (1.6%)	3 (2.4%)	
N classification[a]				0.88
N0	170 (29.9%)	134 (30.2%)	36 (29%)	
N1	71 (12.5%)	56 (12.6%)	15 (12.1%)	
N2	313 (55.1%)	242 (54.5%)	71 (57.3%)	
N3	14 (2.5%)	12 (2.7%)	2 (1.6%)	
Tumor grade				0.75
1	561 (98.8%)	438 (98.6%)	123 (99.2%)	
2	5 (0.9%)	4 (0.9%)	1 (0.8%)	
3	2 (0.4%)	2 (0.5%)	0 (0%)	
Surgery				0.001
No	304 (53.5%)	256 (57.7%)	48 (38.7%)	
Yes	264 (46.5%)	188 (42.3%)	76 (61.3%)	
Chemotherapy				0.05
No	263 (46.3%)	196 (44.1%)	67 (54%)	
Yes	305 (53.7%)	248 (55.9%)	57 (46%)	
Radiation therapy				0.93
No	153 (26.9%)	120 (27%)	33 (26.6%)	
Yes	415 (73.1%)	324 (73%)	91 (73.4%)	
Death				<0.0001
No	434 (76.4%)	390 (87.8%)	44 (35.5%)	
Yes	134 (23.6%)	54 (12.2%)	80 (64.5%)	
Death caused by cancer				<0.0001
No	450 (79.2%)	406 (91.4%)	44 (35.5%)	
Yes	118 (20.8%)	38 (8.6%)	80 (64.5%)	

[a] T and N classifications are based on American Joint Committee on Cancer-6 staging system.

second events. The results are displayed in Supporting Information Table S2.

4 | DISCUSSION

To our knowledge, this is the first study to develop and test algorithms for identifying second events in oropharyngeal cancers using routinely collected administrative data from a publicly funded health system. One of the most important contributions of this study is leveraging the unique patterns of medical encounters that occur in the setting of second events in order to identify them. Moreover, this study also provides a framework for constructing similar algorithms to identify second events in other cancers.

Ricketts et al developed an automated method to estimate the recurrence of head and neck cancers using a sample from a cancer center in the United Kingdom.[5] In their cohort, there were 122 patients with head and neck cancer of whom 51 had oropharyngeal cancer. The authors reported that their method obtained 52.5% sensitivity and 97.6% specificity. However, the method used mainly relied on procedures after primary treatments (radiation and chemotherapy). Other studies explored other data-driven strategies for identifying recurrences in other cancers.[13–17] For example, Earle et al developed an algorithm for identifying relapses of acute myelogenous leukemia, whereas Chubak et al developed one for detecting breast cancer recurrences. However, these algorithms were not specific for oropharyngeal cancers, and the difference in coding practice between United States and Canada also impeded the applications of the previous methods to Canada data. For example, algorithms that Chubak et al developed included the second malignant neoplasm diagnosis code in physician billing data or second breast cancer record in surveillance, epidemiology, and end results data that highlight the existence of second events and contributed to their algorithms. However, the second malignant tumor diagnosis code in our data (hospitalization data, ambulatory care data, and physician billing claims data) did not work in our algorithms. Additionally, the Current Procedural Terminology and HCPCS are specific for US data.[18–20] When compared to prior studies, our universal health system eliminated the influence of insurance status on the patterns of medical encounters,[18–20] and our algorithms used common data elements available from administrative data sources which are probably accessible in most publicly funded health systems. Moreover, we limited the number of the inputted variables in the algorithms to ensure ease of use and the potential for wide application.

Of note, we developed a set of algorithms with different performance measures so that a specific one can be selected depending on the research purpose. The rationale for prioritizing one performance measurement (eg, sensitivity) over others has been described in previous publications.[10,21] In brief, the high-sensitivity algorithm is preferred for studies in which the aim is to identify all patients with a given condition, such as one that is focused on second events' surveillance after primary treatment. Conversely, the high-PPV algorithm is important when the purpose of a study is to ensure that all identified cases are truly positive, such as when one desires to construct a cohort to test the effectiveness of the new treatment modality for oropharyngeal cancer. In contrast, for a study that requires accurate

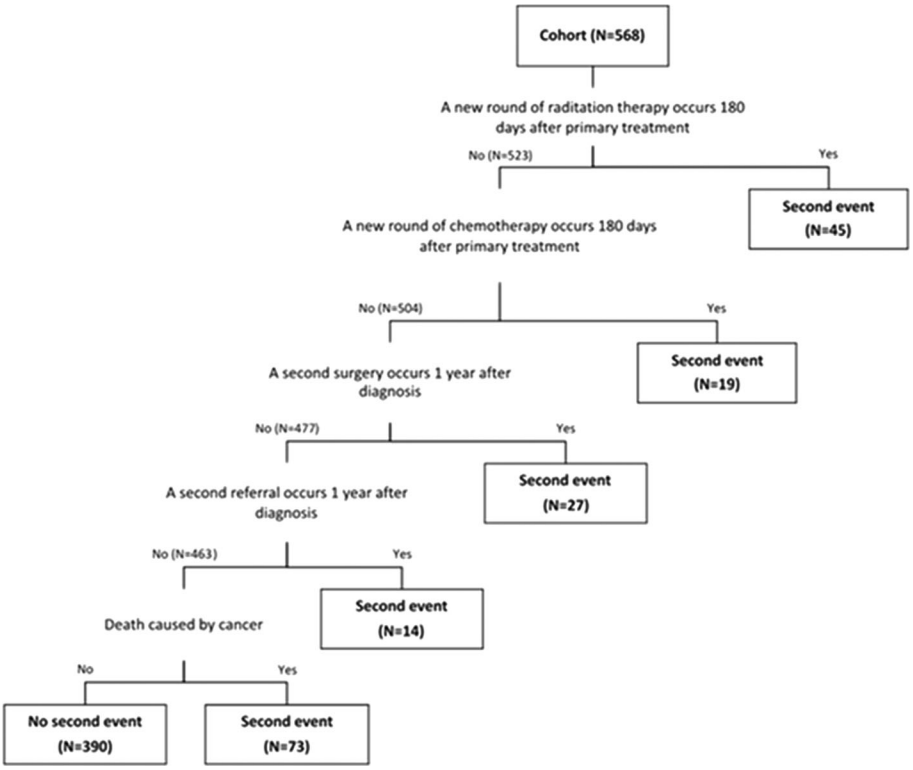

FIGURE 2 Algorithm with high sensitivity of 87.9% (95% confidence interval: 82.2%-93.6%) for identifying second event of oropharyngeal cancer. "Yes" means the criteria was met; "No" means the criteria was not met

identification of true cases and noncases (ie, requiring both high sensitivity and PPV), the combined algorithm that maximizes high sensitivity and high PPV is preferred in which additional chart review can be used to clarify any remaining cases that are uncertain (ie, discordant results from the high-sensitivity and high-PPV algorithms). There are no standards describing accuracy that an algorithm should achieve for identifying second events in an oropharyngeal cancer population. In the literature, many administrative data-based algorithms with greater than 75% sensitivity and/or PPV are considered acceptable (or classified as high) and then widely used in application studies. For example, the hypertension algorithm developed by Quan et al[22] had 75% sensitivity, 94% specificity, and 81% PPV and was deemed as sufficiently sensitive and specific for most research and surveillance purposes.[22–24]

This study has several limitations. First, the study only used data from one region in Canada, so external validation

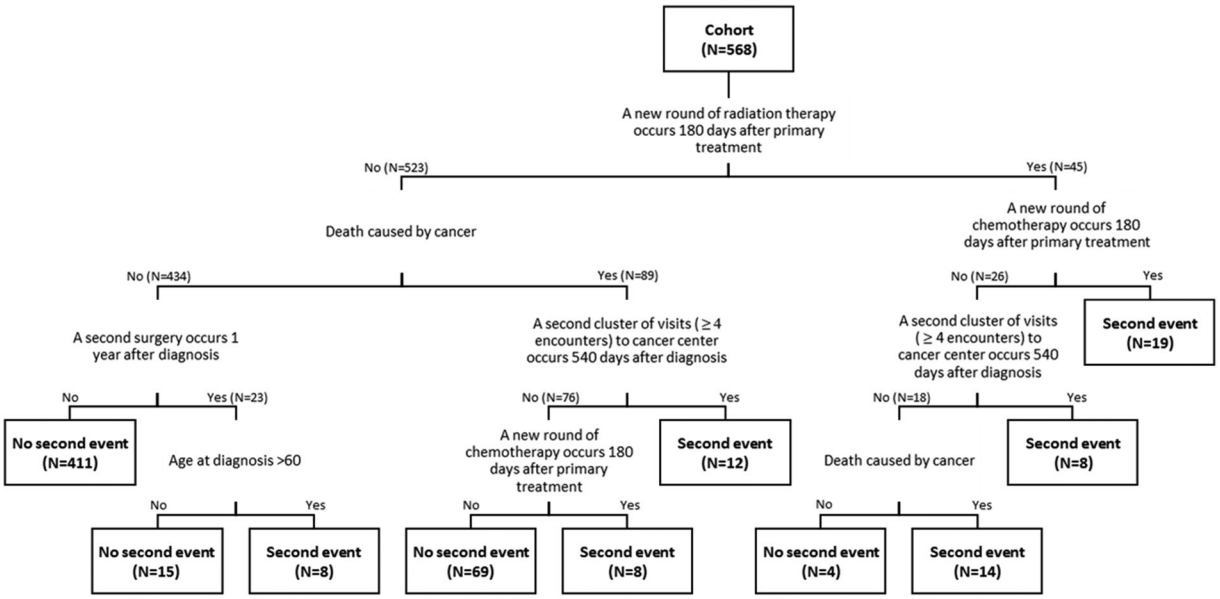

FIGURE 3 Algorithm with high positive predictive value of 94.2% (95% confidence interval: 88.7%-99.7%) for identifying second event of oropharyngeal cancer. "Yes" means the criteria was met; "No" means the criteria was not met

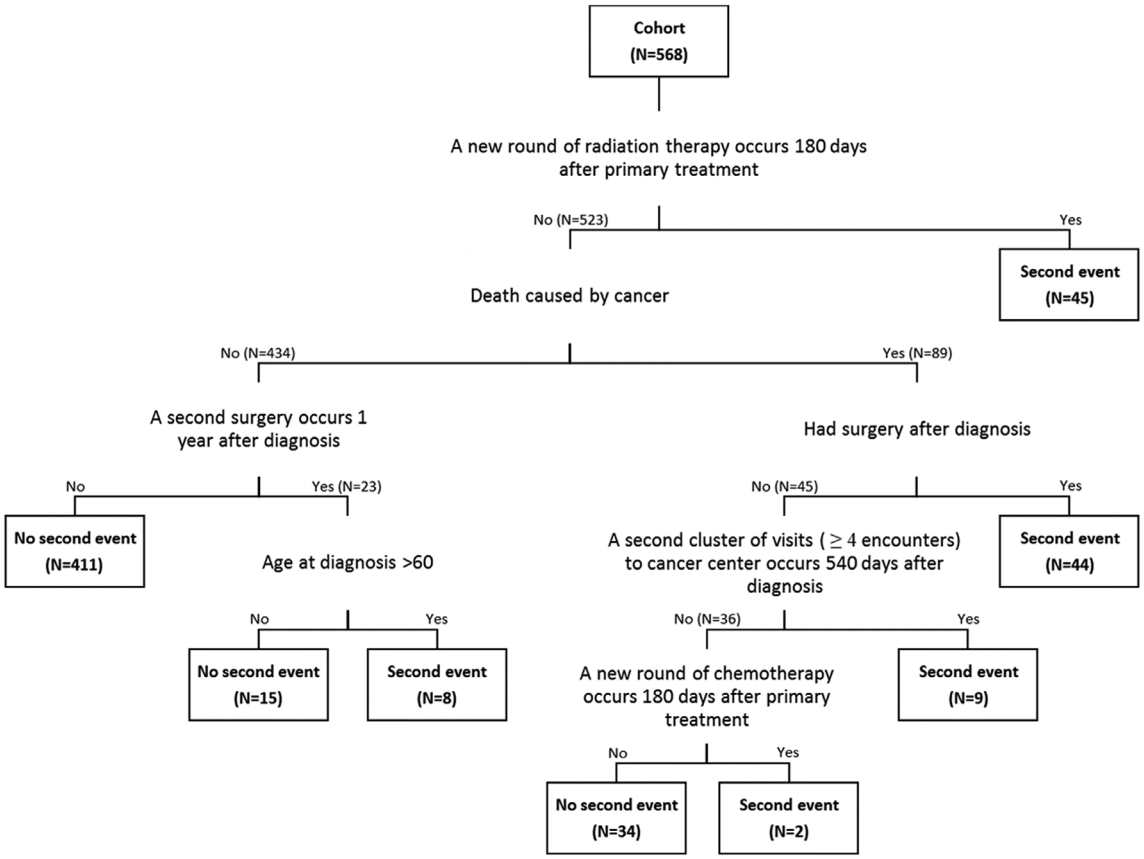

FIGURE 4 Algorithm with balanced sensitivity of 73.4% (95% confidence interval [CI]: 65.6%-81.2%) and positive predictive value of 84.3% (95% CI: 77.4%-91.1%) for identifying second event of oropharyngeal cancer. "Yes" means the criteria was met; "No" means the criteria was not met

is needed by testing the algorithms in other jurisdictions with universal health coverage. Given the similarities across provinces, we do not anticipate significant differences in validity when the algorithms are used in Canada. Also, because the essential part of the proposed algorithm is the pattern of medical encounters of patients with oropharyngeal cancer, we hypothesize that the algorithms may also be generalizable to most patients with oropharyngeal cancer. Second, the rate of second events can impact PPV, so we may observe slightly different PPVs when the algorithms are applied to other populations with different recurrence rates. Similarly, the impact of relatively small number of second events should also be considered when interpreting the validity result of the algorithm. Third, the administrative data did not include the P16 status, which is an important predictor of oropharyngeal squamous carcinoma outcome.[25] However, the purpose of our algorithm was to reliably determine the presence of second events instead of understanding all factors impacting survival. Fourth, our algorithms were not designed to distinguish second primary oropharyngeal cancers from recurrences given their similarity in patterns of care. Finally, improvements of the algorithm are expected, especially for the sensitivity (52.4%) in the high-PPV algorithm. We expect a higher sensitivity if the data algorithm can incorporate the diagnostic imaging tests given that they are frequently used when second events are suspected. Also,

TABLE 3 The validity of the algorithms for identifying recurrences or second primary cancers in the full cohort

Algorithm	Sensitivity (%, 95% CI)	Specificity (%, 95% CI)	PPV (%, 95% CI)	NPV (%, 95% CI)	Accuracy (%, 95% CI)
High sensitivity	87.9 (82.2-93.6)	84.5 (81.1-87.8)	61.2 (54.1-68.4)	96.2 (94.2-98.1)	85.2 (82.3-88.1)
High PPV	52.4 (43.6-61.2)	99.1 (98.2-100.0)	94.2 (88.7-99.7)	88.2 (85.3-91.0)	88.9 (86.3-91.5)
High Accuracy	73.4 (65.6-81.2)	96.2 (94.4-98.0)	84.3 (77.4-91.1)	92.8 (90.5-95.2)	91.2 (88.9-93.5)
Combined method for high sensitivity and high PPV (chart review needed: 109, 19.2%)	87.9 (82.2-93.6)	99.1 (98.2-100.0)	96.5 (93.1-99.9)	96.7 (95.1-98.3)	96.7 (95.2-98.1)

Abbreviations: CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

the timing of second event is also important, especially for cancer-free survival analysis. Further study is needed to develop algorithm to determine the timing of second events.

5 | CONCLUSION

The proposed algorithms achieved acceptable validity for identifying second events in oropharyngeal cancers using widely available administrative data in a universal health system in Canada. Further studies focused on external validation of the algorithms and development of an algorithm to improve performance and to identify timing of recurrence are needed prior to their widespread use.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest with the contents of this article.

ORCID

Yuan Xu https://orcid.org/0000-0003-2100-9437
Steven C. Nakoneshny https://orcid.org/0000-0002-8920-0310
Joseph C. Dort https://orcid.org/0000-0002-0858-187X

REFERENCES

1. Pipkorn P, Licata J, Kallogjeri D, Piccirillo JF. Association of symptoms and clinical findings with anticipated outcomes in patients with recurrent head and neck cancer. *JAMA Otolaryngol Head Neck Surg*. 2018;144(8):738-745.
2. Lowe VJ, Boyd JH, Dunphy FR, et al. Surveillance for recurrent head and neck cancer using positron emission tomography. *J Clin Oncol*. 2000;18(3):651-658.
3. Carrell DS, Halgrim S, Tran DT, et al. Using natural language processing to improve efficiency of manual chart abstraction in research: the case of breast cancer recurrence. *Am J Epidemiol*. 2014;179(6):749-758.
4. Warren JL, Yabroff KR. Challenges and opportunities in measuring cancer recurrence in the United States. *J Natl Cancer Inst*. 2015;107(8):djv134. https://doi.org/10.1093/jnci/djv134.
5. Ricketts K, Williams M, Liu ZW, Gibson A. Automated estimation of disease recurrence in head and neck cancer using routine healthcare data. *Comput Methods Programs Biomed*. 2014;117(3):412-424.
6. LeBlanc MaT R. Combining estimates in regression and classification. *J Am Stat Assoc*. 1996;91:1641-1650.
7. Marczinski CA, Stamates AL, Maloney SF. Differential development of acute tolerance may explain heightened rates of impaired driving after consumption of alcohol mixed with energy drinks versus alcohol alone. *Exp Clin Psychopharmacol*. 2018;26(2):147-155.
8. Frayne SM, Miller DR, Sharkansky EJ, et al. Using administrative data to identify mental illness: what approach is best? *Am J Med Qual*. 2010;25(1):42-50.
9. Borzecki AM, Wong AT, Hickey EC, Ash AS, Berlowitz DR. Identifying hypertension-related comorbidities from administrative data: what's the optimal approach? *Am J Med Qual*. 2004;19(5):201-206.
10. Chubak J, Pocobelli G, Weiss NS. Tradeoffs between accuracy measures for electronic health care data algorithms. *J Clin Epidemiol*. 2012;65(3):343-349.
11. Winkelmayer WC, Schneeweiss S, Mogun H, Patrick AR, Avorn J, Solomon DH. Identification of individuals with CKD from Medicare claims data: a validation study. *Am J Kidney Dis*. 2005;46(2):225-232.
12. Warren JL, Feuer E, Potosky AL, Riley GF, Lynch CF. Use of Medicare hospital and physician data to assess breast cancer incidence. *Med Care*. 1999;37(5):445-456.
13. Chubak J, Yu O, Pocobelli G, et al. Administrative data algorithms to identify second breast cancer events following early-stage invasive breast cancer. *J Natl Cancer Inst*. 2012;104(12):931-940.
14. Hassett MJ, Uno H, Cronin AM, Carroll NM, Hornbrook MC, Ritzwoller D. Detecting lung and colorectal cancer recurrence using structured clinical/-administrative data to enable outcomes research and population health management. *Med Care*. 2017;55(12):e88-e98.
15. Kroenke CH, Chubak J, Johnson L, Castillo A, Weltzien E, Caan BJ. Enhancing breast cancer recurrence algorithms through selective use of medical record data. *J Natl Cancer Inst*. 2015;108(3):djv336. https://doi.org/10.1093/jnci/djv336.
16. Earle CC, Nattinger AB, Potosky AL, et al. Identifying cancer relapse using SEER-Medicare data. *Med Care*. 2002;40(8 suppl):IV-75-81.
17. Wong SL, Ricketts K, Royle G, Williams M, Mendes R. A methodology to extract outcomes from routine healthcare data for patients with locally advanced non-small cell lung cancer. *BMC Health Serv Res*. 2018;18(1):278.
18. Aday LA, Andersen RM. The national profile of access to medical care: where do we stand? *Am J Public Health*. 1984;74(12):1331-1339.
19. Ayanian JZ, Kohler BA, Abe T, Epstein AM. The relation between health insurance coverage and clinical outcomes among women with breast cancer. *N Engl J Med*. 1993;329(5):326-331.
20. Hayward RA, Shapiro MF, Freeman HE, Corey CR. Who gets screened for cervical and breast cancer? Results from a new national survey. *Arch Intern Med*. 1988;148(5):1177-1181.
21. Benchimol EIMD, To T, Griffiths AM, Rabeneck L, Guttmann A. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. *J Clin Epidemiol*. 2011;64(8):8.
22. Quan H, Khan N, Hemmelgarn BR, et al. Validation of a case definition to define hypertension using administrative data. *Hypertension*. 2009;54(6):1423-1428.
23. Pickard SS, Gauvreau K, Gurvitz M, et al. Stroke in adults with coarctation of the aorta: a national population-based study. *J Am Heart Assoc*. 2018;7(11):e009072. https://doi.org/10.1161/JAHA.118.009072.
24. Turin TC, Saad N, Jun M, et al. Lifetime risk of diabetes among First Nations and non-First Nations people. *CMAJ*. 2016;188(16):1147-1153.
25. Ragin CC, Taioli E. Survival of squamous cell carcinoma of the head and neck in relation to human papillomavirus infection: review and meta-analysis. *Int J Cancer*. 2007;121(8):1813-1820.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Xu Y, Kong S, Cheung WY, Quan ML, Nakoneshny SC, Dort JC. Developing case-finding algorithms for second events of oropharyngeal cancer using administrative data: A population-based validation study. *Head & Neck*. 2019;41:2291–2298. https://doi.org/10.1002/hed.25682