CrossMark
← click for updates

Predicting pulmonary tuberculosis in immigrants: a retrospective cohort study

Courtney Heffernan [1], Alexander Doroshenko[1,2,3], Mary Lou Egedahl[1],
James Barrie[4], Ambikaipakan Senthilselvan[3] and Richard Long[1,3]

Affiliations: [1]Tuberculosis Program Evaluation and Research Unit, Dept of Medicine, University of Alberta, Edmonton, AB, Canada. [2]Division of Preventive Medicine, Dept of Medicine, University of Alberta, Edmonton, AB, Canada. [3]School of Public Health, University of Alberta, Edmonton, AB, Canada. [4]Dept of Radiology, University of Alberta, Edmonton, AB, Canada.

Correspondence: Richard Long, University of Alberta, Room 8325, Aberhart Centre, 11402 University Avenue, Edmonton, AB T6G 2J3, Canada. E-mail: richard.long@ualberta.ca

ABSTRACT Our objective was to investigate whether pulmonary tuberculosis (PTB) can be predicted from features of a targeted medical history and basic laboratory investigations in immigrants.

A retrospective cohort of 391 foreign-born adults referred to the Edmonton Tuberculosis Clinic (Edmonton, AB, Canada) was studied using multiple logistic regression analysis to predict PTB. Seven characteristics of disease were used as explanatory variables. Cross-validation assessed performance. Each predictor was tested on two outcomes: "culture-positive" and "smear-positive". Receiver operating characteristic (ROC) curves were generated and the area under the ROC curve (AUC) was quantified.

Symptoms, subacute duration of symptoms, risk factors for reactivation of latent TB infection and anaemia were all associated with a positive culture (adjusted OR 1.79, 2.24, 1.72 and 2.28, respectively; $p<0.05$). Symptoms, inappropriate prescription of broad-spectrum antibiotics and a "typical" chest radiograph were associated with smear-positive PTB (adjusted OR 2.91, 1.55 and 12.34, respectively; $p<0.05$). ROC curve analysis was used to test each model, yielding AUC=0.91 for the outcome "culture-positive" disease and AUC=0.94 for the outcome "smear-positive" disease.

PTB among the foreign-born can be predicted from a targeted medical history and basic laboratory investigations, raising the threshold of suspicion in settings where the disease is relatively rare.

@ERSpublications
In high-income, low tuberculosis incidence countries, certain clinical characteristics should raise the threshold of suspicion to confirm a timely diagnosis http://ow.ly/bRDZ30iPurz

Introduction

Among the G7 (Group of Seven) Nations in 2015 there was a low tuberculosis (TB) incidence ranging from 3.2 per 100 000 population in the USA to 17.1 per 100 000 population in Japan [1]. In the same year, Canada reported the second lowest national TB incidence rate at 4.6 per 100 000 population [2]. These low national incidence rates obscure disparate rates for vulnerable population subgroups. For example, in 2015 the TB incidence for the foreign-born in Canada was 14.8 per 100 000, accounting for 71% of diagnosed cases, and the incidence among Indigenous peoples was 17.1 per 100 000, while the rate for Canadian-born non-Indigenous persons was 0.6 per 100 000 [2]. In the Western Canadian province of Alberta (population 4 146 000), foreign-born persons contributed ~90% of the 211 cases diagnosed with TB in 2015 ([3]; Provincial TB Consultant, Alberta Health Services, Edmonton, AB, Canada). The ability of Canada to meet global targets for TB pre-elimination, *i.e.* 1 case per 100 000 population by 2035, is threatened given the high proportion of foreign-born cases in Canada and the knowledge that, with routine programming, the incidence of TB will not decline by >2% per year once the foreign-born contribute >70% of cases annually [4–6].

An observation important to TB elimination in high-resource, low-incidence settings is that expertise in managing the disease declines relative to the falling burden of disease [7, 8]. In countries like Canada, where universal, transferrable and equitable healthcare is presupposed, but where people can live remotely over vast geographic areas, managing the disease is becoming ever more challenging [7].

A tool that predicts the likelihood of pulmonary TB (PTB) may improve the efficiency of diagnosis by aiding the clinical decision-making process. More-timely PTB diagnoses result in reduced individual morbidity and mortality, and potentially diminish the pool of latently infected individuals by interrupting transmission [9]. Thus, the main objective of this study was to investigate whether PTB among foreign-born persons referred for TB services in Alberta could be predicted from the medical history, results of a complete blood count and a chest radiograph.

Methods

Design

This retrospective cohort study was undertaken to derive predictions using multiple logistic regression for two main outcomes: culture-positive PTB disease (irrespective of smear status) and more infectious, smear-positive, culture-positive disease (hereafter "infectious PTB"). *A posteriori* considered features of disease were used as potential predictors, including the following routinely collected clinical, laboratory and radiographic information: 1) symptoms (respiratory and constitutional), 2) duration of symptoms, 3) relative absence of dyspnoea, 4) failure to respond to broad-spectrum antibiotics, 5) presence of epidemiological risk(s) of infection, 6) presence of moderate- or high-risk factor(s) for reactivation of latent TB infection (LTBI), and 7) complete blood count results. Chest radiograph features typical of PTB (*e.g.* new upper lung zone dominant infiltrate with or without cavitation but no discernible adenopathy) were also considered [9]. We tested the influence of each predictor on the outcomes, *i.e.* PTB and infectious PTB, to test the hypothesis that some variables would have a greater predictive value for each.

Settings and selection of study participants

In Alberta, all TB services are delivered through three public health clinics: one in each of the major metropolitan areas of Calgary and Edmonton, and a Provincial TB Clinic serving clients from rural and remote areas of the province using public health networks and telehealth modalities [7]. Recruitment for this study was from the Edmonton TB Clinic (ETBC). Interpreter use was discretionary.

Study population

Consecutive foreign-born adults (aged >14 years) referred to the ETBC between January 1, 2013 and April 30, 2017 were considered eligible (n=391). This sample size was determined to be adequate to detect a change of $p_{(culture/smear=1)}$ from 0.25 to 0.333 or greater when a continuous predictor was increased 1 SD above the mean (equivalent to OR ⩾1.5) with 80% power and 5% level of significance after allowing all other variables to have an R^2 of 0.15 with the continuous predictor [10].

Within the cohort, the top five countries of origin were the Philippines, India, Somalia, China and Vietnam, accounting for 317 of the 391 (81%) eligible adults.

Methodology

Historical information for each member of the cohort was abstracted from public health records *via* the integrated public health information system retrospectively, and complemented by data gleaned from a standard patient intake form introduced into the ETBC at the outset of the study and by which some data were collected prospectively (see supplementary appendix S1). No members of the cohort had missing

data for any of the explanatory variables. True disease status was defined as having a positive culture for *Mycobacterium tuberculosis* from airway secretions, while infectious PTB cases were determined to be culture-positive from airway secretions and, from that sample, have a positive acid-fast bacilli smear. Culture positivity was determined on the basis of one positive result as per the Canadian TB standards, with every eligible member having at least two sputum specimens collected. Smear positivity was also determined on the basis of one positive result. Remaining members neither grew the organism nor were notified as having disease (n=221).

Each of the predictors was considered related to the outcomes, *i.e.* PTB (culture-positive) and infectious PTB (smear-positive), by some degree. The degree of association was represented by a score from 1 to 4. For example, symptoms ranged from "no symptoms" (score of 1), "one constitutional or one respiratory symptom excluding haemoptysis" (score of 2), "two constitutional, or two respiratory symptoms, or two with one of each, or three constitutional or respiratory symptoms excluding haemoptysis" (score of 3), or "three symptoms with any combination of respiratory and constitutional excluding haemoptysis, and haemoptysis with or without any other symptoms, or four symptoms" (score of 4) [9, 11]. Respiratory and constitutional TB symptoms are generally sensitive but lack specificity and, as such, symptoms alone should not be used for the purposes of TB diagnosis or screening. Dyspnoea is an uncommon symptom of PTB and its absence was considered an independent predictor (see Results).

Predictors were coded, 1–4, and entered into an Excel spreadsheet (Microsoft, Redmond, WA, USA). These data were then imported into Stata version 14 (StataCorp, College Station, TX, USA) for further analysis. Stratification for each predictor is provided in supplementary appendix S2. Chest radiographs were re-read by an independent academic chest radiologist blind to the presence or absence of active disease. The radiologist classified radiograph results as "typical" or "atypical" for adult-type PTB [12].

Statistical analysis

Prediction equations were obtained for both outcomes: PTB or infectious PTB, or not. Multiple logistic regression analysis was used to determine each prediction equation. Collinearity was assessed by using Kendall's τ_b, with none seen. Explanatory variables were treated as continuous, with the exception of sex and chest radiograph, which were dichotomised. Table 1 describes these data by each outcome, *i.e.* PTB (culture-positive) and infectious PTB (smear-positive). Differences in mean age and proportion of males were tested by the two-tailed t-test and the Chi-squared test, respectively. Differences in the mean scores for each clinical feature were assessed using the Mann–Whitney U-test. Purposeful selection was used for model building, such that any predictors significant at $p \geqslant 0.20$ in univariate analysis were included in the final models. Goodness of fit was assessed using the Hosmer–Lemeshow statistic with a nonsignificant p-value indicating a good fit. The multiple logistic regression model was then considered to be the "prediction equation" for the outcome of interest. Receiver operating characteristic (ROC) curves were generated and the area under the ROC curve (AUC) was quantified for both PTB and infectious PTB diagnoses to assess the performance of the multiple logistic regression to correctly predict true disease status.

TABLE 1 Demographic and clinical features by pulmonary tuberculosis (TB) diagnosis status and smear status

	Culture-positive	Culture-negative	p-value	Smear-positive	All others	p-value
Subjects	170	221		74	317	
Age years	46.2±21.0	44.6±16.3	0.42	45.4±21.2	45.3±17.8	0.97
Male	90 (52.9)	111 (50.2)	0.59	41 (55.4)	160 (50.4)	0.44
Clinical characteristics mean score[#]						
Symptoms	2.7±1.14	1.1±0.58	<0.001	3.3±0.87	1.4±0.92	<0.001
Duration of symptoms	3.0±1.24	1.1±0.72	<0.001	3.5±0.89	1.6±1.2	<0.001
Absence of dyspnoea	3.7±0.60	3.9±0.22	<0.001	3.6±0.64	3.9±0.37	<0.001
Antibiotic use	1.9±1.20	1.1±0.53	<0.001	2.4±1.2	1.2±0.75	<0.001
Epidemiological risk of acquiring *Mycobacterium tuberculosis* infection	2.5±0.68	2.8±0.92	0.007	2.3±0.54	2.7±0.87	0.002
Risk factors for reactivation of LTBI	1.4±0.82	1.1±0.51	<0.001	1.5±0.74	1.1±0.66	<0.001
Complete blood count	2.9±0.99	2.1±0.64	<0.001	3.1±1.01	2.3±0.82	<0.001

Data are presented as n, mean±SD or n (%), unless otherwise stated. LTBI: latent TB infection. [#]: refers to the mean of the score, ranging from 1 to 4, characterising the association between the predictor and outcome from clinical experience.

Validation

Model performance was examined by applying the leave-one-out cross-validation method to the full cohort [13, 14]. In this method, each observation is excluded one at a time, with a multiple logistic regression model fitted for the remaining observations; the predicted probability for each outcome (risk score) is then determined using the observed values for the deleted observation. The ROC curve for the initial logistic regression is then compared with the sensitivity and specificity of a 2×2 table generated by cross-validation. The leave-one-out cross-validation method tends to yield a less optimistic prediction, and is routinely used as a test of internal validity and to estimate external validity [14]. All statistical tests were performed using Stata version 14.

Ethics approval

The Health Research Ethics Board panel at University of Alberta approved this study (approval Pro00035012). Alberta Health Services provided administrative and operational approvals. Patients were not involved in the design of this study, but future cohorts are expected to benefit from its results. Informed consent was not obtained because of the retrospective nature of the data used in this study. TB is a time-limited, curable disease so ongoing contact with the clinic is not likely. All data were anonymised and aggregated.

Results

As shown in table 1, distribution was equal by age and sex for the outcomes PTB *versus* culture-negative (no disease) and for infectious PTB *versus* all others (smear-negative, culture-positive disease and culture-negative (no disease)). We observed significant differences in the mean scores of clinical features across disease status.

The final multiple logistic regression model for the outcome PTB tested the seven aforementioned clinical characteristics as potential predictors (table 2). Chest radiograph was not included because people without PTB generally do not have chest radiography results that are "typical" of adult-type PTB. We therefore saw an unstable result for "typical" chest radiograph on PTB. Four predictors were found to have a statistically significant association with the outcome: 1) symptoms, 2) duration of symptoms, 3) risk factors for reactivation of LTBI and 4) complete blood count.

Persons with more symptoms suggestive of TB had a greater odds of having PTB confirmed *via* culture (adjusted OR (aOR) 1.79, 95% CI 1.09–2.96). Moreover, those symptoms appeared to be subacute or chronic in nature ("duration") (aOR 2.24, 95% CI 1.51–3.34). Risk factors for the reactivation of LTBI ("risk factors") also increased the odds of disease (aOR 1.72, 95% CI 1.08–2.72). A chronic disease process was suggested by the association between the results of the complete blood count and a positive *M. tuberculosis* culture (aOR 2.28, 95% CI 1.58–3.31). All other predictors showed a positive relationship with PTB, with the exception of dyspnoea, although they were found not to be statistically significant. The area under the ROC curve (AUC) for culture-positive-only PTB is 0.9135 (figure 1), suggesting a strong combined sensitivity and specificity of our multiple logistic regression model for PTB (disease) *versus* no disease.

Among persons with culture-confirmed disease, having a typical chest radiograph increases the odds of having infectious PTB (aOR 12.34, 95% CI 5.32–28.62). Chest radiograph was included because infectious cases were compared with all others, including smear-negative, culture-positive PTB cases among whom some had "typical" chest radiograph results. Just as in the first multiple logistic regression, the predictor

TABLE 2 Results from the multiple logistic regression model for culture status		
	aOR (95% CI)	**p-value**
Age	0.98 (0.96–1.00)	0.05
Male	1.13 (0.62–2.07)	0.67
Clinical characteristics		
Symptoms	1.79 (1.09–2.96)	0.02
Duration of symptoms	2.24 (1.51–3.34)	<0.001
Absence of dyspnoea	0.85 (0.36–2.00)	0.71
Antibiotic use	1.16 (0.77–1.77)	0.46
Epidemiological risk of acquiring *Mycobacterium tuberculosis* infection	1.06 (0.74–1.51)	0.72
Risk factors for reactivation of LTBI	1.72 (1.08–2.72)	0.02
Complete blood count	2.28 (1.58–3.31)	<0.001

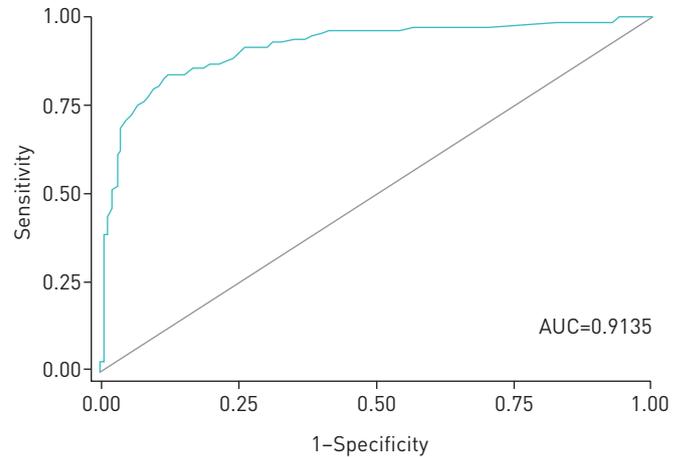aOR: adjusted OR; LTBI: latent tuberculosis infection.

FIGURE 1 Receiver operating characteristic (ROC) curve of the prediction equation for the outcome "culture-positive" (*i.e.* pulmonary tuberculosis). AUC: area under the ROC curve.

"symptoms" was found to have a statistically significant association with the outcome infectious PTB and so too was the predictor "antibiotics" (table 3). The explanatory variable "antibiotics" measured prescription history of broad-spectrum antibiotics in the 6 months preceding the date of diagnosis, with no appreciable or sustained relief of symptoms. Infectious PTB appears to be more strongly predicted than PTB alone (AUC=0.9423) (figure 2).

Compared with the logistic regression model for the outcome PTB, with an AUC of ~0.91, cross-validation yielded a sensitivity of 83.8% and specificity of 87.3% (positive predictive value 91.4%, negative predictive value 77.0% and accuracy 85.1%). Compared with the logistic regression model for the outcome infectious PTB with an AUC of ~0.94, cross-validation yielded a sensitivity of 92.8% and specificity of 73.9% (positive predictive value 94.3%, negative predictive value 68.9% and accuracy 87.7%). Goodness of fit was tested using the Hosmer–Lemeshow statistic and both models were found to have a good fit (p=0.18 for "PTB" and p=0.95 for "infectious PTB"). A nonstatistically significant result indicates failing to reject the null hypothesis that the data do not fully, or properly, describe the outcome of interest.

Discussion

This retrospective cohort study used consecutive referrals of foreign-born adults to the ETBC to determine if predicting who would ultimately have smear-negative, culture-positive and smear-positive, culture-positive disease is possible. Ideally the results of these models can be used to develop a tool to assist in the identification of PTB using clinically available tests and routinely collected information. We suggest a targeted medical history for patients referred to TB services or presenting with general respiratory complaints in primary care.

Respiratory and constitutional symptoms combined with a long duration of symptoms, the results of a complete blood count suggesting anaemia of chronic disease or micronutrient deficiency, with a normal or

TABLE 3 Results from the multiple logistic regression model for infectious pulmonary tuberculosis (smear-positive, culture-positive disease)

	aOR (95% CI)	p-value
Age	1.00 (0.98–1.02)	0.79
Male	1.23 (0.57–265)	0.59
Chest radiograph	12.34 (5.32–28.62)	<0.001
Clinical characteristics		
Symptoms	2.91 (1.80–4.68)	<0.001
Duration of symptoms	0.94 (0.60–1.46)	0.80
Absence of dyspnoea	0.62 (0.30–1.26)	0.18
Antibiotic use	1.55 (1.10–2.18)	0.01
Epidemiological risk of acquiring *Mycobacterium tuberculosis* infection	0.94 (0.52–1.69)	0.84
Risk factors for reactivation of LTBI	1.35 (0.86–2.14)	0.18
Complete blood count	1.45 (0.96–2.18)	0.07

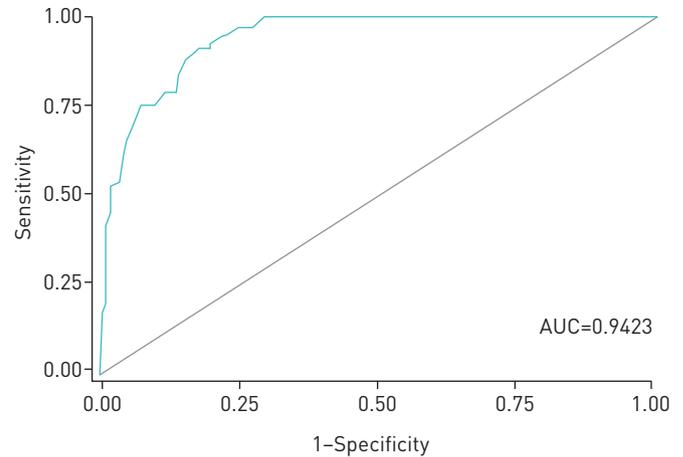aOR: adjusted OR; LTBI: latent tuberculosis infection.

FIGURE 2 Receiver operating characteristic (ROC) curve of the prediction equation for the outcome "smear-positive" (*i.e.* infectious pulmonary tuberculosis). AUC: area under the ROC curve.

low leukocyte count suggesting a process other than community-acquired pneumonia, and risk factors relating to the reactivation of LTBI provide good predictive value of culture-positive PTB [8, 16]. Prescription(s) of broad-spectrum antibiotics and chest radiograph features that are typical of adult-type PTB predict more infectious, smear-positive disease. This constellation of features suggests that smear-positive cases may have had more contact with the healthcare system and that their diagnoses may have been delayed, thereby supporting the implementation of a valid instrument that helps make a more-timely diagnosis. We have not looked into whether these delays are associated with increased transmission, but foreign-born persons in Alberta have previously been reported not to transmit to the same extent as Canadian-born patients [15]. While the number of cases from the top five nationalities reported in this study has increased over time, the vast majority of these cases are likely the result of reactivation of latent infection acquired abroad.

With respect to the relative absence of dyspnoea, all members of the cohort presented similarly. A relative absence of dyspnoea in TB is understood to be due to parallel reductions in ventilation and perfusion in the diseased lung, with the result that gas exchange is relatively well preserved. It is an underappreciated clinical feature of PTB [9, 16]. The fact that dyspnoea appears to be protective in our study group indicates that culture-positive PTB patients are very much like nondiseased persons, with neither group found to be dyspnoeic on examination in general. This feature, therefore, ought to be retained given its ability to offer clinically important information that can distinguish PTB from other respiratory illnesses (*e.g.* relative absence of dyspnoea might support a differential diagnosis in favour of PTB over pneumonia in primary care) [16, 17]. "Risk factors" measured the moderate- and high-risk factors for PTB in persons with LTBI as described in the 7th edition of the Canadian TB Standards [18].

Among migrants, time since arrival in the new country is a classical predictor of progressing to active disease. Two papers published in low-incidence settings (Canada and the USA) suggest that this is likely due to active case-finding activities in persons referred by immigration authorities for medical surveillance post-arrival. If medical surveillance referrals are excluded, the risk of reactivation of LTBI among persons not referred for medical surveillance appears to be the same year over year [19, 20]. As such, it was not included as a predictor in these models, although recent arrivals who were referred for medical surveillance would be indirectly captured by two of the clinical characteristics contributing to the score. These include a high-risk lung scar on chest radiography (high medical risk factor for the reactivation of LTBI) and a past history of TB (part of the epidemiological risk characteristic). Among those not referred, time of arrival in Canada is not captured directly or indirectly.

There has been growing interest in using prediction models for TB. Recent papers have evaluated the significance of predicting a prognosis of PTB, best practices for screening and meta-analyses for ruling out (negative prediction) TB among HIV/AIDS co-infected individuals [21–24]. This is the first study to indicate the value of prediction in a low-incidence setting as a diagnostic aid for PTB among foreign-born persons. Suspicion would appear to be raised relative to the number of presenting clinical features. As such, it is theoretically possible to reach a threshold of suspicion (risk score) after which point sputum is collected, isolation is indicated or an appropriate referral is made. If this is the case, clients may be triaged more appropriately, and loss to follow-up and delays in diagnosis (major contributors to the slow rate of decline in incidence) may be reduced.

Given the high proportion of foreign-born TB cases in Canada, assuming a 2% reduction in the incidence of TB per year is quite generous and, at that pace, our national pre-elimination target is >85 years away if

TB programming does not change drastically [4–6]. While targeted screening efforts and overseas programme improvements are encouraged, technological innovation may assist the timeliness of diagnosis in the potential foreign-born cases already arrived [6, 25]. Our study findings corroborate the literature that, in low-incidence settings where TB is not often considered, patients presenting with infectious disease appear to have multiple missed opportunities to have their diagnosis made.

One limitation of this study is that the data are retrospective; validation with a prospectively collected dataset would strengthen the findings. However, because TB is relatively rare in high-income countries like Canada, collecting a sufficient sample of patients who are later confirmed to, in fact, have disease, prospectively, is logistically prohibitive. Leave-one-out cross-validation is considered to be a sound method of overcoming this limitation with the results of our cross-validation showing strong internal validity [26]. The external validity of these predictive models remains an open question and further testing is required in future studies. Data for all patients were collected, in part, from a form that was introduced into the clinic for the purpose of quality improvement in TB client management and can be considered enhanced surveillance. This eliminated the possibility that we would have missing data for the members of our cohort. In other clinical settings, missing data could present a problem.

Globally, TB is a major cause of mortality with one person dying approximately every 18 s from the disease and its effects [27]. In Canada, the overall incidence is low; however, overrepresentation of cases by the foreign-born is a strong indication that in-country epidemiology reflects the epidemiology of disease in the country of origin among immigrants, refugees and visitors [6, 28]. While supporting TB programming in resource-limited settings remains a cost-effective way of reducing prevalence of infection among foreign-born persons arriving in Canada, data suggest that additional in-country interventions are required to reach elimination targets [29]. The foreign-born dominate the cases referred to the TB clinics in Alberta, but not all persons referred have TB. Identifying high-, medium- and low-risk referrals upon receipt by using a tool underwritten by these prediction models could be considered one such programmatic innovation.

Elimination of TB will be achieved by two mechanisms: 1) reducing the pool of latently infected persons through screening and use of preventive therapy, and 2) interrupting transmission wherever possible. From a public health perspective, identifying smear-positive cases is important because they present with the more communicable form of the disease. Using predictive features specific to the outcome "infectious PTB" provides an opportunity to halt the reproduction of the disease by recommending more-timely treatment, thereby limiting their ability to transmit [30]. This work is recommended to aid a more-timely diagnosis and should not be used as a PTB diagnostic. The gold standard for PTB diagnosis is recovery of the organism in culture, although application of a risk score derived from prediction models may prompt the earlier collection of sputa. As such, we recommend that a targeted medical history include the commonly collected clinical features presented in this work and that they be used in combination to triage cases appropriately in order to more efficiently manage resources.

References

1   World Health Organization. Global Tuberculosis Report. 2016. www.who.int/tb/publications/global_report/en/ Date last accessed: October 12, 2017.
2   Government of Canada. Tuberculosis in Canada 2014: Pre-release. 2014. www.canada.ca/en/public-health/services/publications/diseases-conditions/tuberculosis-canada-2014-pre-release.html Date last accessed: October 12, 2017.
3   Statistics Canada. Population by year, province and territory. www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/demo02a-eng.htm Date last accessed: October 15, 2017. Date last updated: September 27, 2017.
4   World Health Organization. Global Tuberculosis Control: Surveillance, Planning Financing. Geneva, Switzerland, 2008.
5   Long R, Ellis E. TB elimination in Canada: truce or victory? *CMAJ* 2015; 187: 1191–1192.
6   Langlois-Klassen D, Wooldrage K, Manfreda J, *et al.* Piecing the puzzle together: foreign-born tuberculosis in an immigrant-receiving country. *Eur Respir J* 2011; 38: 895–902.
7   Long R, Heffernan C, Gao Z, *et al.* Do "virtual" and "outpatient" public health tuberculosis clinics perform equally well? a program-wide evaluation in Alberta, Canada. *PLoS One* 2015; 10: e0144784.
8   The National Tuberculosis Advisory Committee. Essential components of a tuberculosis control program within Australia. *Commun Dis Intell Q Rep* 2014; 38: E397–E400.

9   Long R. Making a timely diagnosis of pulmonary tuberculosis. *Can Respir J* 2015; 22: 317–321.
10 Hsieh FY, Block DA, Larsen MD. A simple method of sample size calculation for linear and logistic regression. *Stat Med* 1998; 17: 1623–1634.
11 Russell DG. *Mycobacterium tuberculosis* and the intimate discourse of a chronic infection. *Immunol Rev* 2011; 240: 252–268.
12 Lau A, Barrie J, Winter C, *et al*. Chest radiograph patterns and the transmission of tuberculosis: implications for automated systems. *PLoS One* 2016; 11: e0154032.
13 Vach W. Regression Models as a Tool in Medical Research. Boca Raton, CRC Press, 2013.
14 Smith GCS, Seaman SR, Wood A, *et al*. Correcting for optimistic prediction in small data sets. *Am J Epidemiol* 2014; 180: 318–324.
15 Kunimoto D, Sutherland K, Wooldrage A, *et al*. Transmission characteristics of tuberculosis in the foreign-born and the Canadian-born populations of Alberta, Canada. *Int J Tuberc Lung Dis* 2004; 8: 1213–1220.
16 Long R, Maycher B, Dhar A, *et al*. Pulmonary tuberculosis treated with directly observed therapy: serial changes in lung structure and function. *Chest* 1998; 113: 933–943.
17 Bass JB Jr. Dyspnea. *In*: Walker HK, Hall WD, Hurst JW, eds. Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd Edn. Boston, Butterworths, 1990; pp. 78–80.
18 Public Health Agency of Canada, Canadian Lung Association, Canadian Thoracic Society. Canadian Tuberculosis Standards. 7th Edn. Ottawa, Canada, 2013.
19 Walter ND, Painter J, Parker M, *et al*. Persistent latent tuberculosis reactivation risk in United States immigrants. *Am J Respir Crit Care Med* 2014; 189: 88–95.
20 Wang JS, Allen EA, Enarson DA, *et al*. Tuberculosis in recent Asian immigrants to British Columbia, Canada: 1982–1985. *Tubercle* 1991; 72: 277–228.
21 Pefura-Yone, EW, Balkissou AD, Poka-Mayap V, *et al*. Development and validation of a prognostic score during tuberculosis treatment. *BMC Infect Dis* 2017; 17: 251.
22 Watkins RE, Plant AJ. Predicting tuberculosis among migrant groups. *Epidemiol Infect* 2002; 129: 623–628.
23 van't Hoog AH, Onozaki I, Lonnroth K. Choosing algorithms for TB screening: a modelling study to compare yield, predictive value and diagnostic burden. *BMC Infect Dis* 2014; 14: 532.
24 Getahun H, Kittikraisak W, Heilig CM, *et al*. Development of a standardized screening rule for tuberculosis in people living with HIV in resource-constrained settings: individual participant data meta-analysis of observational studies. *PLoS Med* 2011; 8: e1000391.
25 Varughese M, Langlois-Klassen D, Long R, *et al*. Preventing tuberculosis in the foreign-born population of Canada: a mathematical modeling study. *Int J Tuberc Lung Dis* 2014; 18: 405–412.
26 Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. *J Am Stat Assoc* 1982; 78: 316–331.
27 Stop TB Partnership. Stop TB Partnership and its Partners 'Unite to End TB' by Launching a Social Impact Fund and an Accelerator for Impact. 2016. www.stoptb.org/news/stories/2016/ns16_052.asp Date last accessed: October 15, 2017.
28 Jensen M, Lau A, Langlois-Klassen D, *et al*. A population-based study of TB epidemiology and innovative service delivery in Canada. *Int J Tuberc Lung Dis* 2012; 16: 43–49.
29 Schwartzman K, Menzies D. Tuberculosis screening of immigrants to low-prevalence countries. *Am J Respir Crit Care Med* 2000; 161: 780–789.
30 Lohmann EM, Koster BFPJ, le Cessie S, *et al*. Grading of a positive smear and the risk of *Mycobacterium tuberculosis* transmission. *Int J Tuberc Lung Dis* 2012; 16: 1477–1484.