Evaluating the Cancer Risk Management Model (CRMM) – Lung Cancer Module

Authors: Mohsen Sadatsafavi, Zafar Zafari, Craig Mitton, Stirling Bryan

Centre for Clinical Epidemiology and Evaluation, Vancouver Coastal Health Institute, Vancouver, BC

Contact Information: Mohsen Sadatsafavi, MD, PhD
7th Floor, 828 West 10th Avenue
Research Pavilion
Vancouver, BC V5Z 1M9
Tel: 604.875.5178 | Fax: 604.875.5179
Email: msafavi@mail.ubc.ca

Date: 2015/07/31

Contents

Disclaimer

Our group has previously conducted an evaluation of the colorectal cancer module of the Canadian Risk Management Model (CRMM). There are several shared characteristics between the lung cancer and colorectal cancer modules, including reliance on the same modeling framework (ModGEN), the same demographic module (Population Heal Model [POHEM]), and the same conceptual and analytical frameworks. Given this, some parts of the present report are similar to the above-mentioned previous report.

1. Background and objectives

The Canadian Partnership Against Cancer (CPAC), in collaboration with Statistics Canada, has developed the CRMM, a web-enabled computer simulation platform to inform policy and decision making in the Canadian context in various types of cancer. The lung cancer module of the CRMM (CRMM-LC) can be used to inform Canadian guidelines and recommendations for lung cancer screening. Screening for lung cancer has become a focal point of attention with recent evidence, especially from the National Lung cancer Screening Trial (NLST) indicating the ability of low-dose computed tomography (LDCT) in reducing lung cancer mortality(1). However, widespread implementation of LDCT is also costly and therefore the ultimate question is in which subgroups, if any, screening for lung cancer provides the best value for the resources it consumes. Addressing this question requires a framework for quantifying the costs and health outcomes of various screening programs over a sufficiently long time horizon.

The objective of this activity was *to evaluate the lung cancer sub-model of the CRMM (CRMM-LC) with regard to its capacity to evaluate, in terms of cost-effectiveness, decisions regarding lung cancer screening in Canada*. As such, the primary emphasis of the evaluation is on the capacity of CRMM-LC to define screening programs or interventions and perform economic evaluation of such programs in line with established guidelines and best practice recommendations in the field. This is a methodological review of the CRMM-LC. Given the nature of CRMM, evaluating the model structure, internal validity, and capacity to address stakeholders' and consumers' needs is different from evaluating data sources and input parameters. The latter component requires dedicated activity involving cancer epidemiology experts. As such, no explicit numeric results are provided in this report. Rather, qualitative interpretation of the results in terms of their face validity and internal validity (e.g., if input-output relations follow the expected patterns) are provided.

For this evaluation, the team had access to the following components:

- CRMM Web interface (cancerview.ca). Multiple versions of CRMM are available on this Web interface. We used Version 2.1 for this evaluation.

- Management Data workbooks (Lung cancer Management Data workbook and Lung Cancer Screening Module Costing Workbook [the latter is for version 2.2 but was assessed assuming it is the most updated version of input values for CRMM-LC]).
- Peer-reviewed publications and manuscripts related to CRMM and the lung cancer module(2–6).

2. Methodological evaluation of CRMM-LC

In evaluating CRMM-LC, we have undertaken two broad steps of **1)** evaluating the model against the Consolidated Health Economic Evaluation Reporting Standards (CHEERS), a reference checklist for evaluation of the quality of cost-effectiveness analyses(7), **2)** evaluating the face validity of the model in terms of input-output relations. The remainder of the report is structured around these steps. Our concerns, suggestions, and recommendations are highlighted throughout the text and in tabular format at the end.

2.1. Evaluation of CRMM-LC against the CHEERS standard

For evaluation of the CRMM model structure we applied the relevant (methods and results) sections of the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist(7). CHEERS is a checklist for recommended conduct and reporting of economic evaluations. While the present assessment does not consider any specific evaluations, it evaluates the capacity of CRMM-LC to conduct evaluations that are aligned with CHEERS standards. *Table 1* summarizes the results of the implementation this checklist on CRMM model. An itemized description is provided below.

Table 1. Evaluation of the model structure

Model component	Best practice standard		Assessment
Target population	Should validly represent the Canadian population	✓	The inclusion of data from multiple surveys has created a robust and externally valid model that represents the Canadian population
Study perspective	Preference: societal perspective	?	There does not seem to be options for incorporation of productivity loss in calculations
Comparators	All relevant comparators should be evaluated (or the capacity for their evaluation should exist)	✓	This is out of the scope of this work to evaluate comparators, but the model is flexible enough to incorporate a wide range of screening scenarios
Time horizon	Preference: life time	✓	Given that it is an open population, the life time setting is irrelevant. The framework

			enables specification of start and end date of screening as well as how long individuals are followed in the model, providing total flexibility in setting the appropriate time horizon.
Discount rate	Should be based on Canadian guidelines	✓	Flexible in assigning discount rates including separate rates to costs and health outcomes
Currency and price index	Should use a reference costing year	✓	This is largely tackled in the excel sheet preprocessing the cost components. Costing year is 2008 in the current Data Workbooks. Future analyses can use a more up-to-date reference year.
Model structure	Should be logical, plausible, and valid	?	Rigorously designed with input from a wide range of expertise. However, the approach in modeling screening programs is less standard and somewhat non-intuitive.
Analytic methods	Sound statistical analyses and assumption	✓	Rigorous and valid application of statistical methodology whenever required.
Programming codes	Model structure and codes should be made available.	?	Not available to the evaluation team, but based on multi-year work on ModGen and POHEM, as well as the proven face validity of input-output relations, there is not much concern about the programming codes.
Study parameters	The values, ranges, references, and, if used, probability distributions for all parameters. Report reasons or sources for distributions used to represent uncertainty where appropriate		
Disease progression		?	The natural history of lung cancer has not properly been taken into account. Similarly, the impact of screening has been modeled in a non-standard manner.
Resources and costs		✓	High quality work is behind resource use and costs with information from multiple sources as well as expert opinion. The use of Ontario sources for costing cancer outcomes affects the applicability of results to other provinces.
Health outcomes		✓	Incorporates multiple relevant health outcomes such as cancer prevalence and incidence, mortality, life years, and quality adjusted life years (QALY). The use of HUI index for QALY calculation is sound.

Incremental cost and effectiveness	The model should generate estimates of both costs and effectiveness	✔	Reports on multiple costs and health outcomes. Examples include overall costs, costs of screening, false detection rates, cancer incidence, prevalence, and mortality, quality-adjusted life years.
Characterizing uncertainty	Full incorporation of uncertainty in the evaluation	✖	The model does not take into account uncertainty around model parameters and therefore is not capable of probabilistic analysis.
Characterizing heterogeneity	Full representation of the entire subgroups of populations and variables that might affect the result	?	The demographic module of the model is based on multiple surveys and rigorous characteristics of the Canadian population, with nearly complete characterization of heterogeneity among socio-demographic factors. However, the heterogeneity in disease history (cancer progression) has not been modeled.

2.1.1. Target population

The target population for CRMM-LC is various at-risk (for lung cancer) individuals. CRMM simulates, one by one, individuals from birth to death, representing the Canadian population from the past, present, and future (the latter is based on Statistics Canada's projections). Significant work has been undertaken to ensure high degree of external validity and representativeness of the socio-demographic characteristics of the simulated population. Sources include The Canadian Community Health Survey, the National Population Health Survey, the General Social Survey, and the Canadian Health Survey. The socio-demographic module has recently been updated and has performed robustly in external validation studies(8). It includes key variables such as demographics and socio-economic characteristics, smoking status (including the past status and dynamic changes in smoking status as the simulation progresses), as well as radon exposure. CRMM-LC also enables evaluation to be conducted separately within provinces. Again, highly representative sources have been used to simulate province-specific populations. The representativeness of the target population is especially important in screening as the outcomes are not just incremental cost-effectiveness (ICER) of screening versus no screening (as in conventional cost-effectiveness evaluation) but rather, the overall impact on the budget and the health of the population is of concern. ***Overall, robust modeling of the target population is a fundamental strength of this platform for evaluation of screening strategies and gives this platform an exclusive advantage over alternative modeling choices in the evaluation of screening programs.***

2.1.2. Study perspective

Formal economic evaluations require an explicitly defined perspective for the analysis. Perspective can be that of the patient, care provider, third party payer, or the society. The latter is the recommended one, and we feel it is the most appropriate perspective when a national screening program within a publicly funded health care system is considered. The perspective is especially important in deciding which cost components to be included. In its current setup, CRMM seems to have adopted the third-party payer perspective as only direct costs are included in the cost-effectiveness analysis. Considering only direct costs in the cost-effectiveness analyses is recommended by the US Health Panel on Cost-Effectiveness(9). However, cancer has a substantial impact on productivity loss (indirect costs) from the societal perspective. Also, there are potential out-of-pocket costs as well as costs due to waiting times, travel to seek care, and so on, all of which could be considerable. ***The incorporation of out of pocket as well as indirect costs in the model could have added to the utility of CRMM-LC***.

2.1.3. Comparators

In brief, CRMM-LC provides a vast 'decision space' for robust modeling of various lung cancer screening strategies (see Study Parameters below). Another strong aspect of CRRM-LC is the comprehensive modeling of pathways of care; pathways include all available treatment such as surgery, radiotherapy and (neo-adjuvant) chemotherapy, and palliative therapy, and surveillance after treatment. However, the rate of treatment utilization is modeled to match the provincial/national averages. In parallel, the survival rate of lung cancer is modeled from the Canadian cancer registries. This means that the direct impact of treatment on lung cancer outcomes (e.g., impact of radio therapy in terms of relative risk) has not been directly modeled (aside from a hypothetical new treatment). As a result, it might not be possible to consider the joint impact of specific screening and change in cancer treatment guidelines, not being able to capture the potential interaction between the two (e.g., more expensive therapies become less favorable due to stage shift with the implementation of screening). We acknowledge that this is the limitation most likely imposed by the nature of the data available to the investigator team.

2.1.4. Time horizon

CRMM-LC is an open-population (dynamic cohort) platform, meaning that it does not follow a specific cohort of patients (e.g., 55 years old smoker eligible for screening per NLST criteria). Instead, it follows the entire population over a calendar window (e.g., 2015 – 2055). This is a critical advantage in realistic modeling of the impact of screening programs under gradual implementation and sub-optimal adherence.

2.1.5. Discount rate

Different discount rates can be accommodated in CRMM. We have been able to run the evaluations using a wide range of discounting rates (0% to 20%). Within a single run, the CRMM

runs the analyses with multiple discount rates. While this is not a standard practice in economic evaluations, it provides users with additional flexibility to investigate the outcomes of interest at different discount rates.

2.1.6. Currency and price index

The model appropriately adjusts costs for a given reference year and discounting is implemented. By default, costs are calculated annually and started from year 2008 in the model. We are not sure how easily this parameter can change, but overall recommend using more up-to-date reference costing years in future analyses to increase the relevance of evaluations. Detailed cost calculations are made possible in the companion Data Workbook.

2.1.7. Model structure

In brief, CRMM is a Discrete Event Simulation (DES) of hypothetical Canadian individuals from birth to death. DES models the operation of a system as a discrete sequence of events, with individuals as the unit of simulation. Individual-level simulation is the right choice given multitude of risk factors, the presence of interactions (e.g., sex, smoking, and treatment effect), and the need to incorporate 'history' variables (e.g., history of smoking, previous treatments) (10). The conceptual framework of CRMM in general is that risk factors, screening, and treatment influence the outcomes (outputs of the model), which include cancer incidence and death, the costs of screening and treatment, estimates of cost-effectiveness (cost per life-year gained, cost per quality-adjusted life-year gained) and the impact on taxes and transfers.

To complexity of the context, and extensive number of parameters and their interaction, DES is the best simulation tool for modeling the natural history of the disease. The use of DES comes at the cost of lack of familiarity among stakeholders, lack of standardized software, and computational challenges. The choice of modeling platform as well as the implementation of a Web interface (and a Data Workbook) overcomes many of these.

CRMM-LC classifies lung cancers into two major types: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). The former is categorized into 4 stages (I – IV) and the latter into two stages (limited and extensive). The model incorporates various pathways of care. Overall, the lung cancer sub-model of CRMM seems to be based on a detailed, robust, and valid model structure. The pathways of care seem to have been modeled with a reasonable accuracy. A recent peer-reviewed publication on the lung cancer module provides additional description and face validity for the module to be the backbone of further evaluations.

However, *it is not currently obvious how the model represents the progression of the disease from pre-clinical stages to cure/death.* This is not obvious either in the influence diagram in the Data Workbook, nor in the input parameters and descriptions on the Web interface. It appears that the model starts from diagnosed lung cancer, modeling the pathways of care after diagnosis,

and the future trajectory of the disease as `cured` or `relapsed`. This does not incorporate transition from one stage to another, and therefore can cause difficulties in proper modeling of screening programs that require stage-specific detection rate. As mentioned earlier, it appears that CRMM-LC does not model the efficacy of treatments (e.g., radio therapy). Instead, it models overall cancer survival (as observed in the Cancer Registries and other surveys/studies) under a mixed bag of treatments. This does not enable modeling the impact on the results of changes in the pattern of treatment (e.g., using more aggressive treatments). For the evaluation of lung cancer screening strategies, however, this potential limitation is not of much concern unless changes in current treatment patterns are expected in the future.

It appears this particular method of modeling cancer progression is imposed by the nature of the data available to the developer team. A critical source of data was the Canadian Cancer Registry, which includes life trajectories and outcomes of individuals at the time of cancer diagnosis. It means staging information has only been available cross-sectionally at diagnosis time, limiting the development team to estimate transition across stages. Another effect of this limitation is that not much information has been available to model pre-diagnosis timeline of lung cancer, forcing the team to make non-intuitive assumptions about the impact of screening (such as a time period in which sensitivity and specificity of screening is applicable, and so on). A preclinical cancer phase with a mean time of 1.9 years has been considered but it is modeled in relation with the sensitivity and specificity of screening and stage shift (not corresponding to a well-defined tumor status such as clinical staging). A model that simulates biological progression of cancer (e.g., a preclinical tumor of a given size having a certain probability of being detected in LDCT) does not have to assign time interval to sensitivity and specificity of screening tests. Similarly, stage shift could naturally arise from simulating preclinical stages and the impact of screening in earlier diagnosis of lung cancer. The observed stage shifts in NLST could have been calibration targets not model inputs. However, we also acknowledge that the limited data available to the developer team (and the historical 'backward' development from a model targeted at cancer treatment to a model that investigates screening) has imposed certain restriction to the developer team. Overall, while we understand why the team made certain assumptions and design decisions, we ***believe direct modeling of natural history of lung cancer with using observed results from NLST and other studies as calibration target could have resulted in a simpler and a structure that would be easier to understand.***

2.1.8. Analytic methods

Rigorous approaches and justifiable assumptions are made in this regard. The use of a Weibull distribution, for example, that accommodates a non-constant hazard is a valid and well justified choice. The use of two-stage survival analysis for modeling relapse in CRMM provides a valid and robust statistical support.

Models of this level of complexity require dedicated calibration and validation attempts. The demographic module of CRMM-LC has undergone extensive calibration(8). In addition, the screening module of CRMM-LC has undergone extensive calibration and validation against NLST(3). Aside, other important aspects of CRMM-LC have been subject to dedicated calibration and validation attempts. Importantly, the model removes lung and colorectal cancer from background mortality risk (thus avoids double counting) and also models the impact of smoking on mortality from other causes. These valid assumptions can have non-negligible consequences on the results.

Smoking trajectories have been fitted using certain years of national data and have validated against other years, as well as against tobacco manufacturers' data. Lung cancer incidence rates were calibrated to the documented incidence in the Canadian Cancer Registry for the year 2005 and validated for the period 1999 - 2009. Lung cancer mortality has been calibrated based on the Canadian Mortality Database for the year 005. ***Detailed set of calibration and validation exercises for CRMM-LC, making the platform a strong and trustworthy framework for the evaluation of lung cancer screening programs.***

2.1.9. Incremental cost and effectiveness

CRMM provides a platform to compare easily the cost and effectiveness of different user-defined screening scenarios and calculate their incremental cost-effectiveness ratio. In addition, the web interface provides functionalities for comparing related scenarios in terms of differences in inputs and outputs. This provides a user-friendly way of comparing scenarios not just by their costs but also in terms of other model outputs.

2.1.10. Study (model) parameters

CRMM-LC takes a large number of input parameters, typical of sophisticated models of this level of complexity. This will surely enable the analyst/decision maker to have substantial control over evaluation parameters and features. Importantly, several input parameters pertain to structural assumptions (e.g., the way smoking projections are made), thus enabling the analyst to performed structural sensitivity analyses. Other parameters represent the course of lung cancer, impact of screening/treatment, and the performance of the health care technologies and services (including screening). The evaluating team does not see the need for populating multiple parameters as a drawback, rather as the consequence of the complexity of the landscape underlying the disease and decisions that are to be made. Another aspect of CRMM is that it jointly models multiple cancers (lung, lung, cervical, and so on). As many risk factors such as smoking and obesity affect multiple as it will ultimately enable estimating the effect of programs and interventions.

Notable parameters that are likely to affect screening results are as follows:

Smoking: CRMM currently provides multiple parameter inputs for smoking. By default, the model assumes that the recent smoking trends will continue into the future, but provides very flexible options to model various modification of this assumption (e.g., reducing smoking prevalence in a given year or over a range of years, for targeted age range). It is possible to base the evaluation on different predictions regarding future smoking behavior. Due to the impact of smoking on lung cancer, we recommend to the end-users of the model to explicitly perform sensitivity analyses of cancer screening scenarios under different assumptions about smoking.

Radon exposure: detailed information on the radon exposure, divided across major cities as well as provinces, is provided. The model is capable of incorporating scenarios regarding changes in radon exposure (e.g., reducing it to the acceptable levels according to Canadian guidelines).

Cancer incidence and progression: It is governed by three broad set of parameters: a) incidence rate and stage distribution (from the Canadian Cancer Registry), b) risk equation modifiers (modeling the impact of smoking and radon exposure), and c) progression which models advancement of cancer to the next stage (or death). Lung cancer incidence is well characterized and extensively validated, representing the past, current, and future incidence with high degree of reliability.

Screening parameters: CRMM-LC is quite flexible in modeling lung cancer screening scenarios. There are multiple parameters that define a screening program. Overall, the interface is quite flexible in designing a customized screening program. Examples include annual screening, three annual screenings and biennial screening, and so on. The model also accommodates for potentially assigning lower quality of life weights to individuals with a false positive results over a user-defined time. Based on our evaluation of the CRMM (version 2.1), the platform can in general accommodate the following features

1. Eligibility and implementation: enables the analyst to model age criteria for screening (can be province-specific), years for implementation and termination of screening, recruitment attempts, participation rates, gradual uptake (phase-in period), rescreening rate and frequency.
2. Performance of screening: includes sensitivity and specificity, impact of screening on future incidence, false positive rates and outcomes, radiation risk, stage shift due to positive or negative screening, lung cancer sojourn times, and uptake and complications of follow-up procedures.
3. Survival benefit beyond stage shift,
4. Screening costs and costs of follow-up invasive procedures
5. Change in smoking cessation with screening
6. Detailed cost inputs

Overall, these are critical parameters in evaluation of cancer screening programs. The large parameter space will enable the analyst to define customized screening interventions, and the developer team to envision additional calibration attempts to fine-tune the model with the availability of new evidence. The evaluation team was provided with an exemplary list of scenarios as a template for evaluation; we confirm that CRRM-LC accommodates not only the list but many other cancer screening strategies.

Resources and costs: Resource use and costs for all different screening scenarios, treatments, and events are derived from Canadian data and implemented fully in the model. In general, there is quite an amount of flexibility in modeling the cost profile of screening strategies. Although implementing some complicated scenarios (e.g., time-dependent cost profile) will require 'tweaking' by model developers, we find the model to be flexible enough on this dimension. The CRMM makes a tradeoff between the aggregate and detailed cost calculations: The model inputs aggregate cost parameters grouped by types (e.g., diagnostics, drugs, hospitalization). Costs of starting up the program was not considered. This can be a focus of future developments.

Detailed calculation of these costs is performed in the accompanying Data Workbook. This enables the developer team to work with a manageable number of parameters while the end-user has the flexibility of modifying very specific cost values (e.g., unit cost of bone scan). This is a clever tradeoff and a commendable feature of CRMM. ***Overall, costs are modeled comprehensively and flexibly.*** On the other hand, ***the sole reliance on the Ontario data for cost calculations undermines the validity of results for the other provinces***.

Health outcomes: Health outcomes are expressed in terms of quality-adjusted life years (QALYs) and are calculated based on utility values derived from Classification and Measurement System of Functional Health (CLAMES) (11,12). The choice of QALYs is a positive aspect of CRMM but we have not reviewed the robustness of the utility values. Appropriate modification can generate alternative health outcomes (e.g., setting all utility values to 1) will generate estimates of life years gained from screening as a secondary output of the model. The Data Workbook provides an interface that enables modification of the input utility values. The use of health utility index (HUI) as weights for quality of life is a positive feature of the platform, due to the comparability of HU weights across multiple conditions, and the availability of high quality Canadian data on HUI weights.

2.1.11. Characterizing parameter uncertainty

In general, random variation and uncertainty in a simulation model can be categorized into three broad terms(13). Stochastic uncertainty refers to the inevitable uncertainty in outcomes even within a single individual. Stochastic uncertainty should be removed from the analysis in population-based evaluations. Heterogeneity (or first-order uncertainty) refers to the variation in outcomes due to differences in causal factors (e.g., difference in age resulting in difference in

time to metastasis). When making decisions for the whole population is concerned, the effect of heterogeneity also needs to be removed, but decisions can be made more efficient by stratification of decisions across identifiable subgroups(14). Finally, <u>parameter uncertainty</u> (or second-order uncertainty) refers to the uncertainty in our knowledge of the parameters governing the nature of the disease condition and the context in which it occurs (e.g., our uncertainty about the sensitivity of the screening test). Incorporating uncertainty in decision models requires the capacity for probabilistic analysis. Probabilistic analysis entails assigning probability distribution to all uncertain model parameters, and creating multiple runs of the model such that within each iteration, the results are generated based on a set of random draws from the model inputs.

By simulating the outcomes across many (multiple millions) of individuals and averaging the results, CRMM removes the effect of stochastic uncertainty and heterogeneity. By incorporating the capacity to run the simulation is customized fashion for different subgroups of individuals, CRMM enables stratified decision making. However, and unfortunately, ***CRMM is not a probabilistic model and does not capture uncertainty in decision-making for different screening scenarios.*** Full incorporation and reporting of second-order uncertainty in decision analysis is a requirement and a recommendation by major guidelines and best practice standards(15). However, we acknowledge that aside from reporting of formal cost-effectiveness analyses, the vast output of the model, combined with significant degree of freedom in varying the input parameters for deterministic sensitivity analysis provide the end-user with means to quantify the sensitivity of outputs in particular set of input parameters. CRMM-LC therefore provides alternative means for exploring uncertainty in the results, but the current standards for economic evaluation explicitly require the incorporation of probabilistic analysis in the results and we anticipate that this will be recurring issue in the peer review or expert review economic evaluations based on CRRM.

2.1.12. Characterizing heterogeneity
As described above, heterogeneity is well captured in CRMM through generating a representative sample of Canadian population in terms of their sex, age, province of residence, income quintile, and health-related quality of life. However, the model cannot fully incorporate heterogeneity in other aspects. For example, it does not seem that the model is capable of modeling conditional sensitivity and specificity as a function of individual's characteristics. This, nonetheless, seems achievable through further involvement of the development team.

2.2. Face validity of the lung cancer model
We appraised the **face validity** of the CRMM model concentrating on lung cancer by manipulating the key input parameters and investigating if the direction of outcome changes stays in line with

our expectation. This part of a validation assures us that there is a rational relationship between inputs and outputs of the model(16). Unexpected results can indicate programming error or implausible assumptions. Summary of the selected scenarios evaluated in our face validity exercises can be found in **Table 2**.

The model performed robustly in all face validity exercises, with the change in output occurring where expected, in the direction that was expected, and generally within the magnitude that was expected. ***Our detailed evaluation of model inputs, as well as input-output relations has made us confident about the veracity of the underlying structure and implementation***. In addition, the time requirement for running the scenarios were not prohibitive and 'production-level' analyses (e.g., based on tens of millions of simulations) are generally manageable.

Table 2. Brief description of the face validity tests are provided below. All simulated scenarios are based on Monte Carlo simulation of size 1,000,000 – screening scenarios with N=2,000,000

Scenario	Name on *CancerView*	Change in variables	Expected outcome	Results
Zero incidence (for BC only)	MS_Zero_Incidence3	For both sexes and for all ages, set values to 0 (for BC) in Cancer parameters->Lung cancer->Incidence->Experts only->LC calibrated incidence rates	Zero lung cancer incidence, declining lung cancer prevalence, zero lung-cancer death after a wash-in period	Minor warning: occasional lung cancer deaths even in year 2050; non-declining prevalence. Can be due to immigration into the province.
Exaggerated impact of smoking on lung cancer	MS_Smoking_LC_10X	Increase the coefficients by 10 times in Cancer parameters->Lung cancer->Risk equation coefficients ->Smoking coefficient in lung cancer risk equation	Higher incidence, prevalence, and mortality from lung cancer	OK
A near-perfect new treatment	MS_Perfect_New_Tx2	Relative risk of treatment set to 0.01 in Cancer parameters->Lung cancer->Treatment	Very low levels of LC death, high levels of LC prevalence	OK
Utility values set to zero	MS_UTIL_0	All utility values set to 0 in Population Health Parameters->Average health utility of population by age	QALY should be zero	OK
Treatment costs set to zero	MS_TX_$_01	All cost values set to 0 in Population Health Parameters->Average health utility of population by age	Costs should be zero	OK
Screening sensitivity=0, specificity=1	MS_SC_Sn0Sp1	Sensitivity=0, specificity=1 for all columns in Cancer Parameters->Lung cancer->Screening->Early Detection->Sensitivity and specificity of screening modalities	Should mimic no screening outcomes (lower cancer death)	OK

No participation in BC and Ontario	MS_SC_NOBCON	Start and end years set to 9999 for BC and ON in Cancer Parameters->Lung cancer->Screening->Eligibility, recruitment, and participation->Year to start organized screening program	Should mimic no screening in BC and ON	OK
Reverse stage shift	MS_SC_REVSTSHIFT	All stage shifts to stage IV set to 1 Cancer Parameters->Lung cancer->Screening->Costs->Stage Shift->Stage shift due to screening (non-small cell lung cancer)	Should worsen the outcomes of screenings compared to no screening	OK
Screening costs set to 0 (in BC and ON)	MS_SC_$_0	All costs for BC and ON set to 0 in Cancer Parameters->Lung cancer->Screening->Costs->Screening costs	Reduce the overall costs and make costs of screening to be zero	OK
Lower frequency of screening	MS_SC_FREQ_0	Changed to biennial screening (from manual) by modifying Cancer Parameters->Lung cancer->Screening->Eligibility, recruitment, and participation->Frequency of lung cancer screening in organized program	Should lower screening costs but increase cancer mortality	OK

3. Conclusion

CRMM-LC is a state-of-the-art platform representing years of ground work by Statistics Canada, CPAC, and other agencies. Our overall assessment is that ***CRMM provides a unique opportunity to Canadian authorities in making their decisions and recommendations about lung cancer screening objective, transparent, and evidence-informed***. The input from multiple expert teams (statisticians, clinical experts, and policy experts) has resulted in a rigorous evaluation platform. Extensive model calibration and validation has significantly added to the credibility of results. The latter is a major difference between the lung cancer and colorectal cancer module that was previously reviewed by our group. Our examination of the model provides reassuring results about the face validity of the model and its capacity to validly inform lung cancer screening

policies. ***Keeping in mind a few limitations of the platform, we highly recommend utilizing CRMM-LC as a decision tool on formulating evidence-informed recommendations and policies in lung cancer screening***.

CRMM is a micro-simulation model of lung cancer (and other cancers), enabling robust and valid modeling of the complex natural history of cancer, multiple factors affecting the history, and multitude of outcomes that will be of interest to both epidemiologists and decision makers. CRMM cannot only enable evaluation of the lung cancer screening strategies, it can also act as a reference platform for evaluation of other interventions in the pathway of cancer prevention and treatment. We are not currently aware of any other platforms, in Canada or elsewhere, in cancer or other diseases, that provides a comparable level of functionality. Among the limitations that we have encountered, only the lack of capacity for probabilistic analysis is a relatively major one and the one that will require substantial investment in re-designing the platform.

CRMM is equipped with an ***advanced Web interface*** that provides detailed outputs of the analyses, enabling the user to explore not only the basic results informing a cost-effective analysis, but also myriad of additional outputs regarding the epidemiology of the disease as well as indices of health services use. This is also useful to test face validity, sensitivity to assumption and inputs. The companion document (Data Workbook) provides critically important additional information outlining the model structure and detailed calculation of costs and utility values and probabilities.

Currently, a major drawback of this platform for economic evaluation of lung screening strategies is lack of consideration of parameter uncertainty and consequently, lack of capacity for probabilistic analysis. This means the platform will not be able to generate measures of uncertainty (e.g., credible intervals around the outcomes and the incremental cost-effectiveness ratio [ICER], cost-effectiveness plane and acceptability curves) and value of information metrics. Contemporary economic evaluation guidelines strongly require incorporation of probabilistic analysis in evaluations(13). Indeed, in models that the relation between input and output is non-linear, even the calculation of the point estimates of outcomes and ICERs needs to be based on probabilistic analysis(15).

Another important consideration is the structural assumptions involved in modeling lung cancer. The model makes aggregate and independent assumptions on incidence, stage distribution of lung cancer, the rate of relapse, and death. The central role of cancer registry data in informing the model inputs has resulted in lung cancer trajectories starting from 'diagnosed' cancer. This is problematic when modeling cancer screening scenarios. The developer team has done a great job in reconciling such a model with screening evaluations. However, had the model been designed through explicit modeling of cancer progress (from preclinical to various clinical stages to death), screening could have been modeled more intuitively. We do not see the current

approach to be invalid, but we feel any requirement for update (e.g., arrival of new evidence) would force the developer team to calibrate several parameters of the model in a less intuitive process.

CRMM is inevitably a 'black box' in terms of model structure and assumptions, and the complex inner workings of the platform and the input-output structure will require constant and long-term involvement of Statistics Canada in maintenance and upgrading the platform. This should be of little concern given the commitment and support from the agency for this type of work. The complexity of CRMM comes at the cost of the requirement for long-term (and perhaps perpetual) support from the development team. It is imperative that Statistics Canada maintains an up-to-date version of the model, a task that we feel cannot be relegated to CPAC or other agencies. The long-term investment of the developer institution (Statistics Canada) in the ModGen platform(17) hosting CRMM is reassuring that support will be available.

Table 3. Summary of key issues and suggestions.

Issue	Suggestions
Lack of capacity for probabilistic analysis	Further development of the platform to accommodate parameter uncertainty. Use of statistical techniques to reduce computational time and need for nested simulations (see Conclusions)
Model structure does not represent the biological evolution of the disease	This is most likely due to lack of the availability of data. The developer team could, however, use the data and evidence available as calibration targets towards developing a model that properly captures lung cancer progression from pre-clinical stages (e.g., in-situ carcinoma, benign nodules, pre-malignant nodules->stage I->stage II->stage III->stage IV
Inability to perform evaluation from the societal perspective	Incorporating indirect costs (productivity loss)
Requirement for changing the model structure to explore other scenarios, especially customized screening strategies based on, for example, risk or patient characteristics	Continuous cooperation between the developer team and stakeholders.
Incorporation of province-specific costs as well as indirect costs	This does not seem to require major updates in model structure. Given the documented variation in cancer and costs across provinces, we recommend that any evaluation of screening explicitly explores, through sensitivity analyses, the overall impact of known sources of variations between provinces on the results of screening.

References

1. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011 Aug 4;365(5):395–409.

2. Louie AV, Rodrigues GB, Palma DA, Senan S. Measuring the population impact of introducing stereotactic ablative radiotherapy for stage I non-small cell lung cancer in Canada. The Oncologist. 2014 Aug;19(8):880–5.

3. Government of Canada SC. Health Reports – Performance of the Cancer Risk Management Model Lung Cancer Screening Module [Internet]. 2015 [cited 2015 May 21]. Available from: http://www5.statcan.gc.ca/olc-cel/olc.action?objId=82-003-X201500514169&objType=47&lang=en&limit=0

4. Fitzgerald NR, Flanagan WM, Evans WK, Miller AB. Eligibility for low-dose computerized tomography screening among asbestos-exposed individuals. Scand J Work Environ Health. 2015 Jul 1;41(4):407–12.

5. Evans WK, Wolfson MC, Flanagan WM, Shin J, Goffin J, Miller AB, et al. Canadian Cancer Risk Management Model: evaluation of cancer control. Int J Technol Assess Health Care. 2013 Apr;29(2):131–9.

6. Evans WK, Wolfson M, Flanagan WM, Shin J, Goffin JR, Asakawa K, et al. The evaluation of cancer control interventions in lung cancer using the Canadian Cancer Risk Management Model. Lung Cancer Manag. 2012;1(1):25–33.

7. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health J Int Soc Pharmacoeconomics Outcomes Res. 2013 Apr;16(2):231–50.

8. Manuel DG, Tuna M, Hennessy D, Bennett C, Okhmatovskaia A, Finès P, et al. Projections of preventable risks for cardiovascular disease in Canada to 2021: a microsimulation modelling approach. CMAJ Open. 2014 Apr;2(2):E94–101.

9. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. JAMA J Am Med Assoc. 1996 Oct 16;276(15):1253–8.

10. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M, et al. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--2. Value Health J Int Soc Pharmacoeconomics Outcomes Res. 2012 Oct;15(6):804–11.

11. McIntosh CN, Connor Gorber S, Bernier J, Berthelot J-M. Eliciting Canadian population preferences for health states using the Classification and Measurement System of Functional Health (CLAMES). Chronic Dis Can. 2007;28(1-2):29–41.

12. Evans WK, Connor Gorber SK, Spence ST, Will BP. for the Population Health Impact of Disease in Canada (PHI). Health state descriptions for Canadians. Cancers. Statistics Canada, catalogue no. 82–619- MIE2005001. Ottawa: Statistics Canada, 2005.

13. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD, et al. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--6. Value Health J Int Soc Pharmacoeconomics Outcomes Res. 2012 Oct;15(6):835–42.

14. Espinoza MA, Manca A, Claxton K, Sculpher MJ. The Value of Heterogeneity for Cost-Effectiveness Subgroup Analysis: Conceptual Framework and Application. Med Decis Mak Int J Soc Med Decis Mak. 2014 Jun 18;

15. Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health Econ. 2005 Apr;14(4):339–47.

16. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--7. Value Health J Int Soc Pharmacoeconomics Outcomes Res. 2012 Oct;15(6):843–50.

17. Statistics Canada. Modgen (Model generator) [Internet]. [cited 2012 Mar 19]. Available from: http://www.statcan.gc.ca/microsimulation/modgen/modgen-eng.htm

18. Rojnik K, Naversnik K. Gaussian process metamodeling in Bayesian value of information analysis: a case of the complex health economic model for breast cancer screening. Value Health. 2008;11(2):240–50.