Lung Cancer Risk Prediction Model for National Health Risk Assessment Module

Executive Summary

[Adapted from the report by MAHARITA AB RAHMAN]

Authors:
Madam Maharita Abd Rahman
Dr Izzuna Mudla Mohamed Ghazali

Expert Committee:
Dato' Dr. Abdul Razak Muttalif
Dr. Kamaliah Mohamad Noh
Dr. Nik Harlina Roza Binti Nik Kazim
Dr. Nurhayati Mohd Marzuki
Dr. Azura Deniel
Dr. Nor Saleha binti Ibrahim Tamin
En. Azman Mohammed
Pn. Norhaslida Mohd Nor
Datin Dr. Rugayah binti Bakri
Dr. Junainah binti Sabirin

External Reviewer:
Prof. Dato' Dr. Fuad Ismail
Datu Dr Andrew Kiyu
Dr Jamalul Azizi Abdul Rahman

Information Specialist:
Madam Asmirah Redzuan
Madam Norharlina Che Zakaria

For further information please contact:

Health Technology Assessment Section (MaHTAS)
Medical Development Division
Ministry of Health Malaysia
Level 4, Block E1, Precinct 1
Government Office Complex
62590 Putrajaya.

Tel: 603 8883 1246
Fax: 603 8883 1230

Available at the following website:
http://www.moh.gov.my

Background

Lung cancer is the most common cancer worldwide, accounting for 1.3 million deaths annually. According to the National Cancer Registry, 1,865 cases of lung cancer were diagnosed and registered in Peninsular Malaysia in 2007. The age standardised rate (ASR) for male was 14.7 per 100,000 and 5.6 per 100,000 for female. The incidence was more than two-fold higher among males compared to females. The incidence increased with age and in 2007 the peak of age-specific incidence rate was among the 70-75 age groups. Most of the lung cancers were detected late where 60% of the cases were detected at stage IV, the percentage of lung cancer detected at stage I and II was only 12%.

The United States, National Cancer Institute reported that the lung cancer five-year survival rate (16.3%) is lower than many other leading cancer sites, such as the colon (62.2%), breast (90.0%) and prostate (99.9%). The five-year survival rate for lung cancer was 52.6 percent for cases detected when the disease was localized (within the lungs). However, only 15 percent of lung cancer cases were diagnosed at an early stage. Lung cancer with distant metastases (spread to other organs) the five-year survival rate was only 3.5%.

According to the United States, National Lung Screening Trial (NLST), screening programme is suggested for high risk population. One of the early screening methods is through health risk assessment (HRA) tool. This health risk assessment, also known as health risk appraisal, health & well-being assessment or risk prediction model, is an online questionnaire that asks about lung cancer risk factors and it is completely confidential. Commonly the HRA incorporates three key elements: an extended questionnaire, a risk calculation or score, and some form of feedback i.e. face-to-face with a health advisor or an automatic online report.

In Malaysia, currently HRA modules are available for obesity, mental health, diabetes, heart problems, physical activity and smoking habit. Currently there is no risk assessment prediction model for early detection of lung cancer.

Technical Features:

Cancer risk assessment models / health risk assessment tools are statistical models developed for cancer risk prediction and can be divided into two broad categories:

i. To predict the probability of being diagnosed with a particular cancer, and
ii. To predict the likelihood of carrying a gene mutation that predisposes to a particular cancer or set of cancers

Thus, it is supposed to be useful in clinical decision making. According to Memorial Sloan Kettering Cancer Centre, health risk assessment tools helps clinicians and patients to determine the chance that screening will be beneficial.

Policy Question

i. In the Ministry of Health, should a health risk assessment (HRA) module for lung cancer be introduced as one of the strategies in the prevention of lung cancer under the Malaysian National Cancer Control Programme?
ii. If an HRA module (cancer risk prediction model) for lung cancer is to be introduced, which risk prediction model for lung cancer should be adopted / adapted in Malaysia?

Objectives

i. To assess the effectiveness in terms of predictive accuracy of lung cancer risk

assessment/prediction models

ii. To assess the safety, organizational, ethical issues and economic implications related to risk assessment/prediction models for lung cancer

Methods

Studies were identified by searching electronic databases. The following databases were searched through the Ovid interface: MEDLINE(R) In-process and other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to present. Parallel searches were run in PubMed and EMBASE. Appendix 3 shows the detailed search strategies. No limits were applied to the search. The last search was run on 02nd July 2015. Additional articles were identified from reviewing the references of retrieved articles.

Results and Conclusions

A total of 2,431 titles were identified through Ovid interface, Pubmed and references of retrieved articles. A total of 55 abstracts were screened using the inclusion and exclusion criteria. After critical appraisal, only six full text articles were included in this review.

Of these, five articles were related to effectiveness (predictive accuracy) of different risk prediction models for lung cancer. The other study was validation a study related to the risk prediction model assessed. The six studies included comprised three case controlled studies, one cohort study, one randomised controlled trial and two non-randomized controlled trials. No evidence on safety and cost-effectiveness / cost-cost utility analysis was retrieved.

Effectiveness (predictive accuracy)

Five risk prediction models namely: Liverpool Lung Project (LLP) risk prediction model, Korean risk prediction model, Bach risk prediction model, Spitz risk prediction model and COSMOS risk prediction model were assessed.

Performance of prediction model is commonly measured by means of calibration and discrimination. A well-fitted model has Expected/Observed (E/O) ratio close to 1, a lower number underestimates the condition's incidence and a higher number overestimates the incidences. The concordance (c)-statistics measure model discrimination performance which is similar to area under the receiver operating characteristic curve (ROC). A c-statistics of 1.0 indicates perfect discrimination and 0.5 equivalents to no discrimination between people who develop the condition and those who do not.

Performance of Risk Prediction Models

Liverpool Lung Project (LLP) Risk Prediction Model (United Kingdom)

LLP risk prediction model is an individualized risk prediction model for lung cancer. The data used were based on data from a case-control study of lung cancer in Liverpool; the Liverpool Lung Project (LLP). The model estimated the absolute risk of lung cancer for a given individual and included variables that are readily available to primary care clinicians to facilitate the referral of high risk individuals. The area under the curve (AUC) was 0.71 and from the 10-fold cross validation of the LLP risk prediction model produced AUC statistics of 0.70 which indicated that good discrimination between cases and controls.

LLP risk prediction model Validation

Validation study on three case-control studies, showed that the LLP risk prediction

model had modest discrimination in the European Early Lung Cancer (EUELC) data set (AUC, 0.67 [CI 0.64 to 0.69]) and good discrimination in both Harvard (AUC, 0.76 [95% CI 0.75 to 0.78]) and Liverpool Lung Project Population-Based Cohort (LLPC) (AUC, 0.82 [CI 0.80 to 0.85]) data set. The AUC for smoking duration which was the strongest of the risk factors was 0.63, 0.74, and 0.72 in the EUELC, Harvard and LLPC data sets respectively. Besides that the LLP risk prediction model had moderate overall calibration and improved accuracy at higher values of predicted risks

Korean Risk Prediction Model (Korea)

Korean Risk Prediction Model was an individualized risk prediction model for lung cancer in Korean men using population-based cohort data. The model was configured to estimate the absolute risk that an individual will have lung cancer in eight years as well as to identify the significant risk factors for lung cancer. C-statistic for Korean risk prediction model showed excellent discrimination of 0.864, 95% CI 0.860-0.868. If considering only age and smoking variables, the prediction model also showed excellent discrimination (C-statistic of 0.861, 95% CI 0.857-0.865). The performance of the risk prediction model also showed excellent discrimination with C-statistic of 0.87, 95% CI 0.867-0.876 when using external validation dataset.

Spitz Risk Prediction Models (United States)

Spitz risk prediction models are multivariable models that are constructed separately for never smokers, former smokers and current smokers, incorporating into each model variables that exhibit statistically significant main effects. Each model was well calibrated throughout the entire range of probabilities, as indicated by non-statistically significant Hosmer-Lemeshow goodness-of-fit test statistics (0.777 for never smokers, 0.712 for former smokers and 0.688 for current smokers). Meanwhile by looking at AUC statistics obtained from the validation sets, the AUC was low for never smokers and current smokers compared to former smokers. The results of concordance statistics indicated that the models performed reasonably well to discriminate between cases and controls

Bach Risk Prediction Model (United States)

Bach risk prediction model was developed and validated for individual lung cancer risk that can be applied in both clinical and research settings. The authors examined the predicted 10-year lung cancer risk among subjects enrolled in an ongoing CT screening program. The authors assessed the extent of variation in risk among a cohort of individuals who met typical eligibility criteria for cancer prevention studies to determine whether the risk of lung cancer varies and to ascertain the usefulness of the model as an adjunct to clinical research. The Bach risk prediction model was validated at six study sites; the observed rates of lung cancer across the deciles for the held-out site closely matched those that had been predicted by the corresponding model derived from the five included sites. The cross-validated concordance index was 0.72 and the cross-validated calibration plot by risk deciles was consistent with excellent calibration. The risk prediction model only had a cross–validated concordance index of 0.66.

COSMOS (Continuous Observation of Smoking Subjects) Risk Prediction Model (Italy)

COSMOS risk prediction model was based on epidemiologic and clinical risk factors to estimate the probability of individuals in a high-risk population being diagnosed with lung cancer. This model might be useful to stratify individuals and

select those at high risk for inclusion in screening programs. Another aim was to develop a second model based on baseline CT findings in a screened population, combined with epidemiologic and clinical risk factors, to stratify individuals according to the probability of being diagnosed with lung cancer at repeat screening scans. The second model was proposed for use in large scale screening programs to select lower risk patients in whom the interval between screening CTs can be lengthened and at the same time to identify those at higher risk of lung cancer in whom surveillance intensity might be increased or who might benefit from prevention intervention studies. At the end, 162 lung cancers were detected in 18,095 person-years of observation from baseline, giving a lung cancer detection rate of 0.90 per 100 years. The detection rates (per 100 years) were slightly higher in men (0.95) compared to women (0.78) and in current smokers (0.92) than former smokers (0.79). However, both differences were not significant. No validation study was conducted for the model.

Safety

There was no retrievable evidence on safety issue of risk prediction model for lung cancer.

Cost Implications

There was no retrieval evidence on risk prediction model or HRA for lung cancer, however, the potential direct cost implicated on the designing, developing, and testing is about RM75,000 to RM100,000.

Organizational

Any risk prediction models or HRA modules require computer literate user / patient and internet access. Statisticians with management capability, computer analysis and risk modeling skills are also required to manage the dataset and undertake statistical analyses. This plan will also involve physicians, nutritionists, health counselors (psychologists) and physiologists.

Development of risk prediction models require several considerations including research issues, gaps, any priorities, and alternatives needed to advance the field of cancer risk prediction and make specific recommendations for implementations. The model also needs to be continually calibrated and revalidated.

Any uncertainties associated with risk estimates should be addressed and informed particularly when clinical decision has serious consequences especially for those who are at risk. Because of that, the whole plan of the module should include counseling, further diagnosis with physician as well as further management and treatment.

Conclusion

There was fair level of retrievable evidence for risk prediction models for lung cancer. There were five models identified for predicting lung cancer risk. The LLP risk prediction model and Korean risk prediction model were the best models for predicting lung cancer. LLP risk prediction model appeared to have good to excellent discrimination with area under curve (AUC) 0.71. The LLP risk prediction model also has good ability to distinguish persons who will or will not develop lung cancer by using the predicted 5-year absolute risk. The Korean model is the only model that used Asia population (Korean) and has an excellent discrimination with

c-statistic 0.87

For other risk prediction models, although they were well calibrated and validated, they appeared to have modest ability to discriminate between subjects who will be having lung cancer and those who will not, in the study population.

There was no retrievable evidence on safety related to risk prediction model or health risk assessment module for the detection of lung cancer in the population. None of the module mentioned any health problem including psychological impact among subjects involved.

There was no retrievable evidence on economic evaluation of risk prediction model or health risk assessment module for lung cancer, or cost implication involved in developing a new health risk assessment retrieved. The cost involved in validating a model by a prospective cohort validation study could be very costly depending on the number of study participants and years of follow up. However, the potential direct cost implicated to the designing, developing, testing and commissioning of available one risk prediction model of lung cancer is about RM75,000 to RM100,000.

Risk prediction model or health risk assessment module for lung cancer needs continual validation to give meaningful risk estimate and to ensure its capability in the setting it will be used. The complexity to develop and validate the risk prediction model or HRA module is reflected in the necessary local data required. Dedicated research expertise to create a robust risk prediction model with consistent performance is very important.

Recommendation

Health risk assessment (HRA) module / risk prediction model for lung cancer such as Liverpool Lung Project (LLP) risk prediction model and the Korean risk prediction model need to undergo further validation until a well-fitted model with better predictive ability tailored to Malaysia population is established. The model needs continual validation to determine the consistency of its performance. Besides that, the module should only be introduced as part of comprehensive strategies for lung cancer whereby screening, treatment and rehabilitation is available.