Incorporating genomic data into multivariate risk models for lung cancer

To the Editor: We congratulate Chowdhury et al.[1] for their insightful appraisal of the potential benefits of incorporating genomic data in risk-stratified population screening for breast and prostate cancers, titled "Incorporating Genomics Into Breast and Prostate Cancer Screening: Assessing the Implications."[1] We believe that the very same benefits are relevant to computed tomographic (CT) screening of lung cancer, recently shown to significantly reduce lung cancer mortality in a large randomized study.[2] We outline below how these benefits relate to screening for lung cancer and how incorporating single-nucleotide polymorphism–based data in a polygenic risk model has wider utility in personalized screening than just reducing the costs of screening.[2]

Following publication of the National Lung Screening Trial (NLST) results, several respiratory and oncology groups published their lung cancer screening recommendations, the majority adopting NLST-based eligibility criteria. While this is in keeping with an evidence-based approach, it overlooks the fact that eligibility criteria for NLST are based on risk variables with limited predictive utility and low specificity (i.e., age and smoking history only).[2,3] Given the high cost of CT screening and this limitation of current selection/eligibility criteria, there remain concerns about poor cost effectiveness, high harm-to-benefit ratio, significant overdiagnosis, and high false-positive rates. As a result, there has been considerable debate about how to best implement lung cancer screening into clinical practice, in particular who best to target for screening.

In a recent article, Tammemägi et al.[3] showed that their multivariate, Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO)-based risk model for lung cancer, which included several nongenetic variables (age, smoking history, history of chronic obstructive pulmonary disease (COPD), and body mass index) and "genetic" variables (family history of lung cancer), was more predictive of lung cancer than the NLST-based criteria. Of note, using this model to preselect smokers for screening allowed the majority of the lung cancer cases to be identified by screening only half of those eligible for screening based on NLST selection criteria—roughly doubling the lung cancer detection rate. This is not surprising because the nongenetic variables in the PLCO model are also risk variables indicative of a predisposition to COPD, a disease we and others have shown to be the single best predictor of lung cancer among chronic (older) smokers.[2,3] Recent screening trials have shown that a disposition to COPD, as measured by spirometry (airflow limitation) or CT (emphysema), is found in up to 80% of all lung cancer cases with widely varying age and smoking exposure. Moreover, we recently showed that the lung cancer detection rate per annum in the Pittsburgh Screening Study was fourfold to fivefold higher in smokers with preexisting COPD as compared with those with normal lungs. We suggest that risk-based interventions such as screening for lung cancer include COPD-related variables to best identify the 10–15% of smokers who will develop lung cancer.

With this in mind, we used data from recently published genome-wide association studies to identify the most robust genetic markers for COPD and/or lung cancer. Using these markers, we have developed a polygenic risk model for lung cancer that includes nongenetic and genetic risk variables. The latter combines a family history of lung cancer with single-nucleotide polymorphisms associated with risk of lung cancer and/or COPD in genome-wide association studies. The genetic risk score derived from the combined single-nucleotide polymorphism data is inexpensive relative to one CT scan and adds useful discriminatory data.[2] Like the PLCO model, this model has superior risk predictability over using age and smoking history alone, particularly for "screening out" lower-risk smokers for whom the costs and potential harm may not justify one or two decades of yearly CT screening.[2] This gene-based approach was especially useful for identifying younger and lighter smokers, who would not otherwise be eligible for screening under NLST eligibility despite high risk of lung cancer. This is very relevant to the preselection issue, and the overall sensitivity of screening, as we have estimated that over 50% of all lung cancer cases fall outside the widely recommended NLST-based criteria.[2] We propose that multivariate and polygenic based risk models for lung cancer can improve specificity of screening, thereby reducing costs, harm-to-benefit ratio, and false-positive findings (in absolute terms). Our polygenic model can also improve sensitivity of lung cancer screening by identifying smokers at high risk who fall outside the NLST criteria, accounting for much of the 50% of lung cancers not currently eligible.

Another distinguishing aspect of incorporating genetic data in a risk model for lung cancer is its potential utility in smoking cessation and screening adherence.[2] In contrast to breast and prostate cancers, in lung cancer, the single most relevant risk factor of smoking is readily identified. Although quitting smoking reduces the risk of lung cancer, this benefit dramatically diminishes the later in life a smoker quits; hence 40% of lung cancers occur in ex-smokers. Smoking cessation is integral to most lung cancer screening recommendations because of the benefits of quitting, primarily to minimize lung cancer recurrence and other ("competing") causes of

death, notably coronary artery disease and other smoking-related cancers. There is growing evidence that genetic information promotes positive lifestyle changes in some people, including quitting smoking. We found in a pilot study using our gene-based test that 6-month abstinence rates in smokers with high as compared with moderate risk were twofold greater (40 vs. 20%, respectively).[2] Gene-based testing also increased the uptake of smoking cessation products, suggesting that the risk stratification with a genetic component may help motivate some smokers to attempt quitting.[2] We also have pilot data suggesting that the gene-based risk testing may improve interest and adherence to lung cancer screening.[2] Although these findings require confirmation, they suggest that there are important additional benefits to gene-based risk testing in lung cancer screening and the wider context of reducing deaths from smoking.

We believe that there is growing evidence that incorporating genomic data into multivariate risk models for lung cancer can improve preselection for screening and outcomes from screening for lung cancer.

DISCLOSURE

R.P.Y. and his research, have been supported by grants from the University of Auckland, Health Research Council of New Zealand, and Synergenz BioSciences. Synergenz BioScience holds the patents for gene-based risk tools for lung cancer and chronic obstructive pulmonary disease. R.J.H. declares no conflict of interest.

Robert P. Young, PhD, MD[1] and Raewyn J. Hopkins, BN, MPH[2]

[1]School of Biological Science, University of Auckland, Auckland, New Zealand; [2]Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand. Correspondence: Robert P. Young (roberty@adhb.govt.nz)

REFERENCES

1. Chowdhury S, Dent T, Pashayan N, et al. Incorporating genomics into breast and prostate cancer screening: assessing the implications. *Genet Med* 2013;15:423–432.
2. Young RP, Hopkins RJ, Gamble GD. Clinical applications of gene-based risk prediction for lung cancer and the central role of chronic obstructive pulmonary disease. *Front Genet* 2013;3:210.
3. Tammemägi MC, Katki HA, Hocking WG, et al. Selection criteria for lung-cancer screening. *N Engl J Med* 2013;368:728–736.

doi:10.1038/gim.2013.89