CENTRE DE RECHERCHE
INSTITUT UNIVERSITAIRE
DE CARDIOLOGIE
ET DE PNEUMOLOGIE
DE QUÉBEC
UNIVERSITÉ LAVAL

September 25th, 2024

Review Committee
Geoffrey Ogram Memorial Research Grant
Lung Cancer Canada

Re: Letter of Intent

Dear Committee Members,

I am please to submit a research proposal for the Geoffrey Ogram Memorial Research Grant Program. The title of the project is "Identification of individuals at high risk of lung cancer at the tail end of the polygenic risk score distribution".

For the last two decades, important progress has been made in the field of genetics of lung cancer. Here, we propose that our knowledge of the genetic factors underlying lung cancer has reached a point where a clinically useful polygenic risk score (PRS) can be generated for early detection and prevention. Such development would have broad implications throughout the full lung cancer care continuum.

Thank you for your time and commitment to evaluate grants as part of this program. I look forward to your response.

Sincerely yours,

Yohan Bossé, Ph.D.
Professor, Université Laval
Department of Molecular Medicine
Canada Research Chair in Genomics of Heart and Lung Diseases
Institut universitaire de cardiologie et de pneumologie de Québec – Université Laval
Tel: 418-656-8711 ext. 3725
email: yohan.bosse@criucpq.ulaval.ca
yohanbosselab.com
ORCID: 0000-0002-3067-3711
@YohanBosse

A summary of the proposed research (maximum of 2 pages)

Background

Low-dose computed tomography (LDCT) screening has been shown to reduce lung cancer mortality by 21% in high-risk individuals[1]. Screening eligibility criteria for selecting high-risk individuals are based on two main approaches. First, using categorical age and smoking exposures. For example, the expanded screening eligibility criteria from the US Preventive Services Task Force recommend annual LDCT in adults aged 50 to 80 years who currently smoke or have smoked within the past 15 years with an accumulated 20 pack-years or greater[2]. Second, using risk prediction models based on a number of lung cancer predictors, e.g. the PLCOm2012 model[3]. Unfortunately, with both approaches, some individuals that will be developing lung cancer will never meet current screening eligibility criteria. This is especially relevant in individuals who never smoked.

Projections in the United States indicate that lung cancer deaths related to smoking will continue to decrease in the coming years following the decline of tobacco use[4]. In contrast, the proportion of lung cancer among individuals who never smoked continues to rise[4,5]. The recent expert consensus from the International Association for the Study of Lung cancer has recently established nine priorities for lung cancer screening programs in the next years[6]. One of them is to "establish evidence-based criteria to identify individuals who have never smoked but are at high risk of developing lung cancer, and who would benefit from screening".

In this study, we hypothesized that our knowledge of the genetic factors underlying lung cancer has reached a point where we can use this information to improve risk prediction and delineate individuals at high genetic risk of lung cancer, including individuals who never smoked or those not meeting lung cancer screening eligibility criteria. As part of large international research groups, we have identified genetic markers robustly associated with lung cancer[7-9]. Individually, the lung cancer-associated genetic variants have small effect sizes and thus limited clinical utility. Collectively, however, several genetic variants can potentially be grouped into a polygenic risk score (PRS) to delineate a subgroup of individuals at higher genetic risk of lung cancer (**Figure 1**)[10].

Preliminary results

We have recently evaluated the clinical utility of a lung cancer PRS in a new case-control dataset consisting of 4,002 lung cancer cases from the LORD project and 20,010 ethnically matched controls from CARTaGENE (eBioMedicine 2024)[11]. To do so, we have generated the first genome-wide PRS leveraging the full genome-wide association study (GWAS) data to quantify genetic predisposition. This genome-wide PRS includes more than 1.1 million genetic variants and outperformed all other previously published lung cancer PRS (**Figure 2**). We then demonstrated in LORD/CARTaGENE that the PRS was associated with lung cancer in a dose-response relationship (**Figure 3**). In addition, we also noted that the effect sizes of the PRS were more pronounced in the extreme tails of the PRS, i.e. the odds of lung cancer at higher PRS thresholds representing the >80th, >90th, >95th and >99th percentiles were respectively evaluated at 2.19, 2.72, 3.36, and 4.88. We may thus have a smaller fraction of individuals at the tail end of the PRS distribution (PRS >99th percentile tested in this study) with an odds ratio in the same range as *BRCA1* (OR = 10.57) and *BRCA2* (OR = 5.85) mutations in the field of breast cancer[12]. In the field of heart diseases, it was demonstrated that some individuals carry a polygenic inheritance with a risk effect similar to mutations causing monogenic forms of the disease[13,14]. Whether this is also observed in lung cancer remains unknown and represents the main goal of this project.

General objective

Identify individuals at high genetic risk of lung cancer at the tail end of the PRS distribution.

Research plan

This project will be performed using two existing datasets, namely LORD/CARTaGENE and UK Biobank. A simplified overview of planned analyses is described below. Support is requested to perform data analyses and report results into a peer-reviewed scientific manuscript.

LORD/CARTaGENE

LORD/CARTaGENE is a retrospective case-control dataset. LORD consists of a single site case-only cohort based on a 20-year data collection (1999 to 2021) of patients who underwent lung cancer surgery at our Institute. CARTaGENE is a population-based project including residents from the province of Quebec enrolled between the ages of 40 and 69 (www.cartagene.qc.ca). Cases from LORD and controls from CARTaGENE have been ethnically matched using ancestry-based principal components into a 1:5 ratio leading to a final dataset of 4,002 cases and 20,010 controls[11].

Cross-sectional analyses in LORD/CARTaGENE

For risk factors that are quantitative in nature, it is useful to define a clinically relevant threshold. However, for a PRS in the field of lung cancer, the thresholds to maximize discrimination of high and low genetic risk are unknown (**Figure 1**). In addition, we do not know if a smaller fraction of individuals at the end-tail of the PRS distribution carries a polygenic inheritance with an effect on risk similar to mutations causing monogenic forms of the disease. Analyses will be performed to identify the upper PRS limit to delineate the high genetic group within smoking groups. A PRS threshold establishing a genetic risk category associated with >5-fold increased risk of lung cancer will be determined and compared across smoking groups. In addition, we will refine our previous genome-wide PRS, which is based on the effects of common variants, with other low-frequency pathogenic variants associated with lung cancer[15]. To identify the most informative PRS thresholds, we will use a multivariate statistical method called Principal Component Analysis with Optimal Scaling (PCAOS)[16]. This method will allow us to gauge the value of the PRS in relation to other conventional demographic, clinical and pathological variables used to define lung cancer. Although exploratory in nature, results from these analyses will be validated in an independent cohort, namely UK Biobank.

UK Biobank

UK Biobank is an open access population-based cohort of nearly 500,000 participants enrolled between the ages of 40 and 69 and prospectively evaluated for a range of health-related outcomes[17]. The definition of lung cancer in this study relies on International Classification of Diseases (ICD) codes, ICD9 and ICD10, and self-reported related codes described previously[18]. Excluding individuals with missing genotypes or phenotypes information as well as those that failed genotyping quality control filters, we have identified 5,419 lung cancer cases and 403,003 controls at baseline. For UK Biobank, incident lung cancer events were also ascertained through national cancer registries. During the median follow-up of 7.2 years, 2,025 incident lung cancer cases were diagnosed.

Cross-sectional and longitudinal analyses in UK Biobank

Retrospective analyses in UK Biobank will be performed as described above to validate the results in LORD/CARTaGENE. Because of the prospective follow-up data in UK Biobank, it is possible to compare lung cancer risk with time across the genetic groups defined by the PRS. Cox proportional hazard models will be used to assess the association between the PRS and lung cancer incidence. The main endpoint will be the incidence of lung cancer during follow-up. Accordingly, lung cancer cases identified at baseline will be excluded from longitudinal analyses. Hazard ratios with 95% confidence intervals will be obtained after adjustment for age, sex, smoking status and the top ancestry-based principal components. Other risk factors included in the PLCOm2012 model[3] will also be considered as covariates. We will calculate absolute risk based on the difference in lung cancer incidence over a 5-year period between genetic risk groups defined based on the most informative PRS threshold that will be identified in LORD/CARTaGENE and validated in the retrospective part of UK Biobank.

Outcome

Advances in genetics of lung cancer provide a new solution to quantify genetic predisposition captured by the PRS. Results of this project will allow us to understand the consequences of carrying a large number of common risk variants for lung cancer, i.e. polygenic inheritance. This is needed to establish the upper and lower PRS limits to delineate the low and high genetic subgroups for population screening as well as to identify individuals with risk equivalent to large effect monogenic mutations.

References

1. Bonney, A., Malouf, R., Marchal, C., Manners, D., Fong, K.M., Marshall, H.M., Irving, L.B. & Manser, R. Impact of low-dose computed tomography (LDCT) screening on lung cancer-related mortality. *Cochrane Database Syst Rev* 8, CD013829 (2022).
2. Force, U.S.P.S.T., Krist, A.H., Davidson, K.W., Mangione, C.M., Barry, M.J., Cabana, M., Caughey, A.B., Davis, E.M., Donahue, K.E., Doubeni, C.A., Kubik, M., Landefeld, C.S., Li, L., Ogedegbe, G., Owens, D.K., Pbert, L., Silverstein, M., Stevermer, J., Tseng, C.W. & Wong, J.B. Screening for Lung Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA* 325, 962-970 (2021).
3. Tammemagi, M.C., Katki, H.A., Hocking, W.G., Church, T.R., Caporaso, N., Kvale, P.A., Chaturvedi, A.K., Silvestri, G.A., Riley, T.L., Commins, J. & Berg, C.D. Selection criteria for lung-cancer screening. *N Engl J Med* 368, 728-36 (2013).
4. Jeon, J., Holford, T.R., Levy, D.T., Feuer, E.J., Cao, P., Tam, J., Clarke, L., Clarke, J., Kong, C.Y. & Meza, R. Smoking and Lung Cancer Mortality in the United States From 2015 to 2065: A Comparative Modeling Approach. *Ann Intern Med* 169, 684-693 (2018).
5. Tseng, C.H., Tsuang, B.J., Chiang, C.J., Ku, K.C., Tseng, J.S., Yang, T.Y., Hsu, K.H., Chen, K.C., Yu, S.L., Lee, W.C., Liu, T.W., Chan, C.C. & Chang, G.C. The Relationship Between Air Pollution and Lung Cancer in Nonsmokers in Taiwan. *J Thorac Oncol* 14, 784-792 (2019).
6. Lam, S., Bai, C., Baldwin, D.R., Chen, Y., Connolly, C., de Koning, H., Heuvelmans, M.A., Hu, P., Kazerooni, E.A., Lancaster, H.L., Langs, G., McWilliams, A., Osarogiagbon, R.U., Oudkerk, M., Peters, M., Robbins, H.A., Sahar, L., Smith, R.A., Triphuridet, N. & Field, J. Current and Future Perspectives on Computed Tomography Screening for Lung Cancer: A Roadmap From 2023 to 2027 From the International Association for the Study of Lung Cancer. *J Thorac Oncol* 19, 36-51 (2024).
7. McKay, J.D., Hung, R.J., Han, Y., Zong, X., Carreras-Torres, R., Christiani, D.C., Caporaso, N.E., Johansson, M., Xiao, X., Li, Y., Byun, J., Dunning, A., Pooley, K.A., Qian, D.C., Ji, X., Liu, G., Timofeeva, M.N., Bojesen, S.E., Wu, X., Le Marchand, L., Albanes, D., Bickeboller, H., Aldrich, M.C., Bush, W.S., Tardon, A., Rennert, G., Teare, M.D., Field, J.K., Kiemeney, L.A., Lazarus, P., Haugen, A., Lam, S., Schabath, M.B., Andrew, A.S., Shen, H., Hong, Y.C., Yuan, J.M., Bertazzi, P.A., Pesatori, A.C., Ye, Y., Diao, N., Su, L., Zhang, R., Brhane, Y., Leighl, N., Johansen, J.S., Mellemgaard, A., Saliba, W., Haiman, C.A., Wilkens, L.R., Fernandez-Somoano, A., Fernandez-Tardon, G., van der Heijden, H.F.M., Kim, J.H., Dai, J., Hu, Z., Davies, M.P.A., Marcus, M.W., Brunnstrom, H., Manjer, J., Melander, O., Muller, D.C., Overvad, K., Trichopoulou, A., Tumino, R., Doherty, J.A., Barnett, M.P., Chen, C., Goodman, G.E., Cox, A., Taylor, F., Woll, P., Bruske, I., Wichmann, H.E., Manz, J., Muley, T.R., Risch, A., Rosenberger, A., Grankvist, K., Johansson, M., Shepherd, F.A., Tsao, M.S., Arnold, S.M., Haura, E.B., Bolca, C., Holcatova, I., Janout, V., Kontic, M., Lissowska, J., Mukeria, A., Ognjanovic, S., Orlowski, T.M., Scelo, G., Swiatkowska, B., Zaridze, D., Bakke, P., Skaug, V., Zienolddiny, S., Duell, E.J., Butler, L.M., Koh, W.P., Gao, Y.T., Houlston, R.S., McLaughlin, J., Stevens, V.L., Joubert, P., Lamontagne, M., Nickle, D.C., Obeidat, M., Timens, W., Zhu, B., Song, L., Kachuri, L., Artigas, M.S., Tobin, M.D., Wain, L.V., SpiroMeta, C., Rafnar, T., Thorgeirsson, T.E., Reginsson, G.W., Stefansson, K., Hancock, D.B., Bierut, L.J., Spitz, M.R., Gaddis, N.C., Lutz, S.M., Gu, F., Johnson, E.O., Kamal, A., Pikielny, C., Zhu, D., Lindstroem, S., Jiang, X., Tyndale, R.F., Chenevix-Trench, G., Beesley, J., **Bossé, Y**., Chanock, S., Brennan, P., Landi, M.T. & Amos, C.I. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat Genet* 49, 1126-1132 (2017).
8. Byun, J., Han, Y., Li, Y., Xia, J., Long, E., Choi, J., Xiao, X., Zhu, M., Zhou, W., Sun, R., **Bossé, Y**., Song, Z., Schwartz, A., Lusk, C., Rafnar, T., Stefansson, K., Zhang, T., Zhao, W.,

Pettit, R.W., Liu, Y., Li, X., Zhou, H., Walsh, K.M., Gorlov, I., Gorlova, O., Zhu, D., Rosenberg, S.M., Pinney, S., Bailey-Wilson, J.E., Mandal, D., de Andrade, M., Gaba, C., Willey, J.C., You, M., Anderson, M., Wiencke, J.K., Albanes, D., Lam, S., Tardon, A., Chen, C., Goodman, G., Bojeson, S., Brenner, H., Landi, M.T., Chanock, S.J., Johansson, M., Muley, T., Risch, A., Wichmann, H.E., Bickeboller, H., Christiani, D.C., Rennert, G., Arnold, S., Field, J.K., Shete, S., Le Marchand, L., Melander, O., Brunnstrom, H., Liu, G., Andrew, A.S., Kiemeney, L.A., Shen, H., Zienolddiny, S., Grankvist, K., Johansson, M., Caporaso, N., Cox, A., Hong, Y.C., Yuan, J.M., Lazarus, P., Schabath, M.B., Aldrich, M.C., Patel, A., Lan, Q., Rothman, N., Taylor, F., Kachuri, L., Witte, J.S., Sakoda, L.C., Spitz, M., Brennan, P., Lin, X., McKay, J., Hung, R.J. & Amos, C.I. Cross-ancestry genome-wide meta-analysis of 61,047 cases and 947,237 controls identifies new susceptibility loci contributing to lung cancer. *Nat Genet* 54, 1167-1177 (2022).

9. **Bossé, Y**. & Amos, C.I. A Decade of GWAS Results in Lung Cancer. *Cancer Epidemiol Biomarkers Prev* 27, 363-379 (2018).
10. **Bossé, Y**. & Martel, S. Germline variants invited to lung cancer screening. *Lancet Respir Med* 7, 832-833 (2019).
11. Boumtje, V., Manikpurage, H.D., Li, Z., Gaudreault, N., Armero, V.S., Boudreau, D.K., Renaut, S., Henry, C., Racine, C., Eslami, A., Bougeard, S., Vigneau, E., Morissette, M., Arsenault, B.J., Labbe, C., Laliberte, A.S., Martel, S., Maltais, F., Couture, C., Desmeules, P., Mathieu, P., Theriault, S., Joubert, P. & **Bossé, Y**. Polygenic inheritance and its interplay with smoking history in predicting lung cancer diagnosis: a French-Canadian case-control cohort. *EBioMedicine* 106, 105234 (2024).
12. Breast Cancer Association Consortium. Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N Engl J Med* 384, 428-439 (2021).
13. Thériault, S., Lali, R., Chong, M., Velianou, J.L., Natarajan, M.K. & Pare, G. Polygenic Contribution in Individuals With Early-Onset Coronary Artery Disease. *Circ Genom Precis Med* 11, e001849 (2018).
14. Ghouse, J., Tragante, V., Ahlberg, G., Rand, S.A., Jespersen, J.B., Leinoe, E.B., Vissing, C.R., Trudso, L., Jonsdottir, I., Banasik, K., Brunak, S., Ostrowski, S.R., Pedersen, O.B., Sorensen, E., Erikstrup, C., Bruun, M.T., Nielsen, K.R., Kober, L., Christensen, A.H., Iversen, K., Jones, D., Knowlton, K.U., Nadauld, L., Halldorsson, G.H., Ferkingstad, E., Olafsson, I., Gretarsdottir, S., Onundarson, P.T., Sulem, P., Thorsteinsdottir, U., Thorgeirsson, G., Gudbjartsson, D.F., Stefansson, K., Holm, H., Olesen, M.S. & Bundgaard, H. Genome-wide meta-analysis identifies 93 risk loci and enables risk prediction equivalent to monogenic forms of venous thromboembolism. *Nat Genet* 55, 399-409 (2023).
15. Panagiotou, E., Vathiotis, I.A., Makrythanasis, P., Hirsch, F., Sen, T. & Syrigos, K. Biological and therapeutic implications of the cancer-related germline mutation landscape in lung cancer. *Lancet Respir Med* (2024).
16. Linting, M., Meulman, J.J., Groenen, P.J. & van der Koojj, A.J. Nonlinear principal components analysis: introduction and application. *Psychol Methods* 12, 336-58 (2007).
17. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T. & Collins, R. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 12, e1001779 (2015).
18. Larsson, S.C., Carter, P., Kar, S., Vithayathil, M., Mason, A.M., Michaelsson, K. & Burgess, S. Smoking, alcohol consumption, and cancer: A mendelian randomisation study in UK Biobank and international genetic consortia participants. *PLoS Med* 17, e1003178 (2020).
19. Dai, J., Lv, J., Zhu, M., Wang, Y., Qin, N., Ma, H., He, Y.Q., Zhang, R., Tan, W., Fan, J., Wang, T., Zheng, H., Sun, Q., Wang, L., Huang, M., Ge, Z., Yu, C., Guo, Y., Wang, T.M.,

Wang, J., Xu, L., Wu, W., Chen, L., Bian, Z., Walters, R., Millwood, I.Y., Li, X.Z., Wang, X., Hung, R.J., Christiani, D.C., Chen, H., Wang, M., Wang, C., Jiang, Y., Chen, K., Chen, Z., Jin, G., Wu, T., Lin, D., Hu, Z., Amos, C.I., Wu, C., Wei, Q., Jia, W.H., Li, L. & Shen, H. Identification of risk loci and a polygenic risk score for lung cancer: a large-scale prospective cohort study in Chinese populations. *Lancet Respir Med* 7, 881-891 (2019).

20. Shi, Z., Yu, H., Wu, Y., Lin, X., Bao, Q., Jia, H., Perschon, C., Duggan, D., Helfand, B.T., Zheng, S.L. & Xu, J. Systematic evaluation of cancer-specific genetic risk score for 11 types of cancer in The Cancer Genome Atlas and Electronic Medical Records and Genomics cohorts. *Cancer Med* 8, 3196-3205 (2019).
21. Graff, R.E., Cavazos, T.B., Thai, K.K., Kachuri, L., Rashkin, S.R., Hoffman, J.D., Alexeeff, S.E., Blatchins, M., Meyers, T.J., Leong, L., Tai, C.G., Emami, N.C., Corley, D.A., Kushi, L.H., Ziv, E., Van Den Eeden, S.K., Jorgenson, E., Hoffmann, T.J., Habel, L.A., Witte, J.S. & Sakoda, L.C. Cross-cancer evaluation of polygenic risk scores for 16 cancer types in two large cohorts. *Nat Commun* 12, 970 (2021).
22. Jia, G., Lu, Y., Wen, W., Long, J., Liu, Y., Tao, R., Li, B., Denny, J.C., Shu, X.O. & Zheng, W. Evaluating the Utility of Polygenic Risk Scores in Identifying High-Risk Individuals for Eight Common Cancers. *JNCI Cancer Spectr* 4, pkaa021 (2020).
23. Fritsche, L.G., Patil, S., Beesley, L.J., VandeHaar, P., Salvatore, M., Ma, Y., Peng, R.B., Taliun, D., Zhou, X. & Mukherjee, B. Cancer PRSweb: An Online Repository with Polygenic Risk Scores for Major Cancer Traits and Their Evaluation in Two Independent Biobanks. *Am J Hum Genet* 107, 815-836 (2020).
24. Hung, R.J., Warkentin, M.T., Brhane, Y., Chatterjee, N., Christiani, D.C., Landi, M.T., Caporaso, N.E., Liu, G., Johansson, M., Albanes, D., Marchand, L.L., Tardon, A., Rennert, G., Bojesen, S.E., Chen, C., Field, J.K., Kiemeney, L.A., Lazarus, P., Zienolddiny, S., Lam, S., Andrew, A.S., Arnold, S.M., Aldrich, M.C., Bickeboller, H., Risch, A., Schabath, M.B., McKay, J.D., Brennan, P. & Amos, C.I. Assessing Lung Cancer Absolute Risk Trajectory Based on a Polygenic Risk Model. *Cancer Res* 81, 1607-1615 (2021).
25. Zhang, P., Chen, P.L., Li, Z.H., Zhang, A., Zhang, X.R., Zhang, Y.J., Liu, D. & Mao, C. Association of smoking and polygenic risk with the incidence of lung cancer: a prospective cohort study. *Br J Cancer* 126, 1637-1646 (2022).
26. Zhu, M., Lv, J., Huang, Y., Ma, H., Li, N., Wei, X., Ji, M., Ma, Z., Song, C., Wang, C., Dai, J., Tan, F., Guo, Y., Walters, R., Millwood, I.Y., Hung, R.J., Christiani, D.C., Yu, C., Jin, G., Chen, Z., Wei, Q., Amos, C.I., Hu, Z., Li, L. & Shen, H. Ethnic differences of genetic risk and smoking in lung cancer: two prospective cohort studies. *Int J Epidemiol* 52, 1815-1825 (2023).

Supporting figures and tables

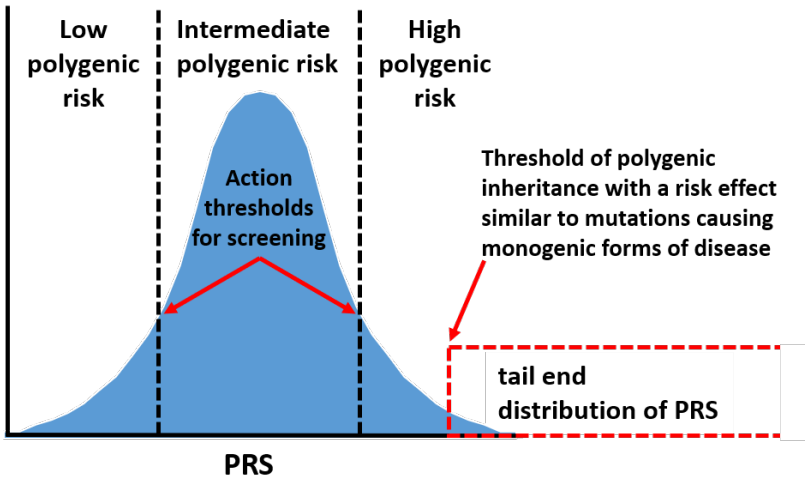

Figure 1. Concepts underpinning this grant application. The blue area represents the distribution of the PRS within the population. Individuals can be classified at low, intermediate and high genetic risk of lung cancer based on specific PRS thresholds that remain to be identified. A more extreme threshold at the tail end of the PRS distribution may also exist to identify individuals with a lung cancer risk at the same extent as *BRCA1*/*BRCA2* mutations in the field of breast cancer.

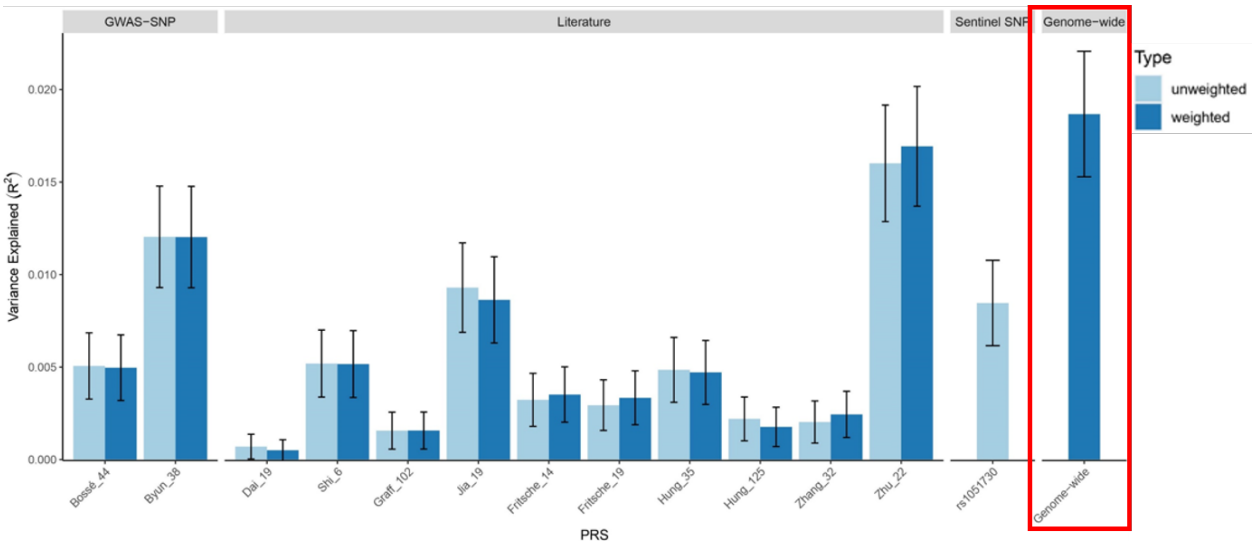

Figure 2. Benchmarking our genome-wide PRS with all previously published PRS in the field of lung cancer in LORD/CARTaGENE. Variance explained by the PRS along with 95% confidence intervals. The set of PRS includes: Two PRS developed in our study[11] from GWAS-nominated loci obtained from Bossé & Amos[9] and Byun et al.[8]; Ten PRS from the literature, named Dai_19[19], Shi_6[20], Graff_102[21], Jia_19[22], Fritsche_14 and Fritsche_19[23], Hung_35 and Hung_125[24], Zhang_32[25], and Zhu_22[26]; A PRS derived from a single sentinel variant (rs1051730) on 15q25; Our genome-wide PRS (red rectangle). For PRS from the literature, two versions were generated: 1) unweighted, representing the sum of risk alleles, and 2) weighted, representing the sum of risk alleles weighted based on their effect sizes [log(OR)]. The number of genetic variants in PRS is indicated in their names.

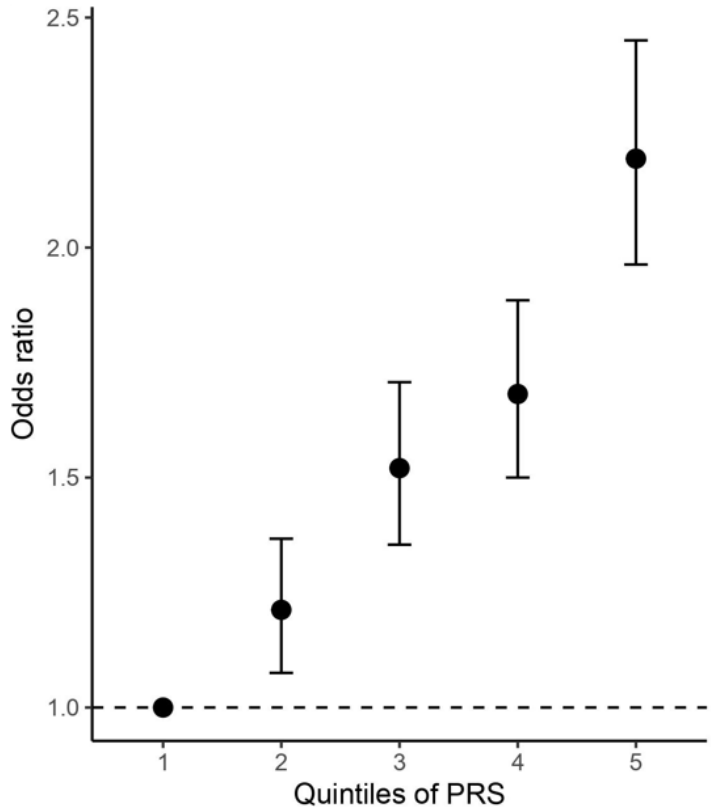

Figure 3. Association of the genome-wide PRS with lung cancer in LORD/CARTaGENE. Odds ratio of lung cancer per quintile increase in the PRS along with 95% confidence intervals. Quintile 1 as the reference including 20% of individuals with the lowest PRS.

Impact statement

Advances in genetics of lung cancer provide a new solution to quantify genetic predisposition captured by the PRS. Results of this project will allow us to understand the consequences of carrying a large number of risk variants for lung cancer, i.e. polygenic inheritance. This is needed to establish the upper and lower PRS limits to delineate the low and high genetic subgroups for population screening as well as to identify individuals with risk equivalent to large effect monogenic mutations. From the identification of high-risk individuals through PRS analyses, we foresee major advancements in lung cancer research and potential downstream clinical utilities.

1) Explain, at least partly, susceptibility/resistance to smoking-related risk of lung cancer through polygenic inheritance captured by the PRS.
2) Contribute to our knowledge on lung cancer development in individuals who never smoked.
3) Inform tobacco control efforts and eventually offer targeted approaches in the high genetic risk group.
4) Provide a solution for cost-effective screening implementation among never-smoking populations.
5) Identify high-risk individuals that do not meet current screening criteria.
6) Find individuals with genetic risk equivalent to mutations such as *BRCA1*/*BRCA2* in breast cancer.
7) Optimize lung cancer screening to delineate true and false positive lung cancer cases, improve the management of indeterminate pulmonary nodules, and inform optimal timing for low-dose CT screening.

Obviously, specific studies for these intended purposes will be needed, but the proposed research project will leverage the discoveries of previous GWAS on lung cancer and translate the new genomic knowledge into results that are one step closer to clinical applications.

A public, non-scientific summary (up to 500 words)

Lung cancer screening by imaging using a special type of computed tomography (CT) scan in high-risk individuals who have ever smoked has been firmly established to reduce lung cancer mortality. However, current screening eligibility criteria rely predominantly on smoking exposure. Worldwide, the number and proportion of lung cancers in people who have never smoked is on the rise. Unfortunately, lung cancer in people who have never smoked will be missed by current screening program. As indicated by a recent expert consensus, one of the main priorities for lung cancer screening in the next years is to "establish evidence-based criteria to identify individuals who have never smoked but are at high risk of developing lung cancer".

In this study, we hypothesized that our knowledge of the genetic factors underlying lung cancer has reached a point where we can identify individuals at high genetic risk of lung cancer. In fact, recent studies conducted by large international research groups have achieved an important step to understand the genetic component of lung cancer by identifying genetic markers robustly associated with the disease. Based on this progress, we can now, for each individual, quantity genetic predisposition to lung cancer by measuring genetic markers and calculating a polygenic risk score (PRS).

We have recently generated a new case-control dataset consisting of 4,002 lung cancer cases and 20,010 ethnically matched controls. In this cohort, the PRS was an independent predictor of lung cancer beyond conventional clinical risk factors including age, sex and smoking. Interestingly, the risk of lung cancer increased with genetic predisposition captured by the PRS. Higher PRS was associated with higher risk of lung cancer. For example, the risk of lung cancer increased by 5-fold in the top percentile of the PRS distribution. We may thus have a small fraction of individuals (approximately 1%) at the tail end of the PRS distribution with a lung cancer risk increase equivalent to the effect of *BRCA1*/*BRCA2* mutations in the field of breast cancer.

The goal of this research project is to identify individuals with high genetic risk of lung cancer regardless of their smoking exposure. Analyses in our case/control cohort will be performed to identify the upper PRS limit to delineate the high genetic subgroup for population screening. Analyses will also be performed stratified by smoking status. Finally, validation will be conducted in an independent cohort, namely UK Biobank, consisting of 5,419 lung cancer cases and 403,003 controls.

This research is essential to identify high-risk individuals that do not meet current screening criteria. Quantifying genetic predisposition with a PRS may provide a cost-effective screening solution for non-smokers. Studying the clinical value of the PRS also has broader implications for prevention and early detection. In contrast to most lung cancer risk factors, genetic markers can be obtained from birth and used early to identify high-risk individuals. Early identification of high-risk individuals is likely to become important for a disease that can be in large part avoided by preventing smoking initiation.

Budget – including justification for supplies, equipment, and personnel associated with the research project. This must include the number of personnel required to complete the work, a description of their experience and/or education level, and their commitment to the project. If your project is part of a larger initiative, please list the other sources of funding. However, the portion of the project proposed in this grant should not be covered by other sources of funding.

Funds requested will be used to support sophisticated bioinformatics and statistical analyses in existing datasets as well as to prepare a scientific manuscript.

Research Staff
Amount: $21,000
Description: A part-time (25% effort) bioinformatician will be hired for this project. Zhonglin Li will fulfill this position. He will be in charge of data stewardship and bioinformatics. The salary is based on the collective agreement of research assistant at U Laval.
Part-time bioinformatician = $21,000 for one year

Knowledge Translation
Amount: $4,000
Description: We request costs for a publication of $3,000. The research manuscript will be freely accessible online. A research allowance of $1,000 is also requested to present the results of this project at a Canadian conference with a lung cancer focus. This will partly cover for registration fees, airfares, hotels, meals, and taxi fares. The main investigator and the graduate student will be travelling.

Total for the project = $25,000.00

CENTRE DE RECHERCHE
INSTITUT UNIVERSITAIRE
DE CARDIOLOGIE
ET DE PNEUMOLOGIE
DE QUÉBEC
UNIVERSITÉ LAVAL

September 24th, 2024

Review Committee
Geoffrey Ogram Memorial Research Grant
Lung Cancer Canada

Re: A statement of support to confirm that the proposed research is feasible

Dear Committee Members,

I am very enthusiastic to provide this statement of support for Dr. Yohan Bossé's research application for the Geoffrey Ogram Memorial Research Grant Program entitled "Identification of individuals at high risk of lung cancer at the tail end of the polygenic risk score distribution". I write this letter in my capacity as a Full Professor in the Department of Medicine at U Laval and Research Director of the *Institut universitaire de cardiologie et de pneumologie de Québec – Université Laval*.

I am confirming that the project proposed by Dr. Bossé is feasible at our Institute. Dr. Bossé is a well-established investigator with a proven track-records in the field of genetics of lung cancer. He is leading an exceptional team with the programming and statistical analysis skills to make data analyses and interpretation realistically achievable. In fact, I see no other team in Canada with this scientific background, expertise and resources to conduct this world class research with this budget. The project is timely, represents a logical extension of his previous work (Boumtje et al. eBioMedicine 2024, PMID 38970920) and leverages knowledge gained over many years in genetics of lung cancer. I foresee enlightening results to assess the clinical utility of polygenic risk score in the care of our patients with lung cancer.

I am thus offering my strongest possible support on this project.

Sincerely,

Mathieu Laplante
Director of Research
Institut universitaire de cardiologie et de pneumologie de Québec – Université Laval

2725, chemin Sainte-Foy, Québec (Québec) G1V 4G5 CANADA
TÉLÉPHONE 418 656-8711 TÉLÉCOPIEUR 418 656-4866
www.CRIUCPQ.ulaval.ca