|  |  |
| --- | --- |
| I. Project Title | |
| **The association of genetically predicted blood protein biomarkers and lung cancer survival**  pCode: XXX | **Date:**  **Feb 4, 2019** |
| **II. Project Group Investigators** | |
| **Project leaders:** Mulong Du, David C. Christiani, Paul Brennan, James Mckay  **Investigators:** ILCCO contributing study/ data PI’s; Mattias Johansson | |
| III. Background | |
| Currently, proteins in blood as biomarkers have been reported to be potentially associated with lung cancer risk and survival. However, the expense of protein detection limited the wide use of proteomics in a large population-based study, which thus contributed to the inconsistent findings from previous studies. Because of the random assortment of alleles transferred from parents to offspring during gamete formation, an approach that uses genetic variants associated with blood protein levels as an instrument to evaluate the effect of genetically predicted protein levels on lung cancer risk and survival could produce a robust result. | |
| IV. Specific Aims | |
| To assess the association of circulating proteins with lung cancer survival. | |
| **V. Methods** | |
| 1. The inverse variance weighting (IVW) method using summary statistics of lung cancer GWAS was utilized to estimate the association. This strategy has been reported in previous studies (*Wu et al., Cancer Res. 2019 Sep 15;79(18):4592-4598.* *Shu et al., Int J Cancer. 2019 Jul 2.*). 2. Sun and colleagues identified 1,927 genetic associations with 1,478 proteins at a stringent significance level, which was used to construct the instrumental variables for assessing associations of protein levels with lung cancer survival (*Sun et al., Nature. 2018 Jun;558(7708):73-79.*). | |
| **VI. Materials or variables needed from the study PIs** | |
| 1. Demographic and clinical descriptions of study samples   For survival:   1. Number of lung cancer cases (need the largest sample size in European). 2. Age, gender, race, smoking status (never/ever/current), package-year of smoking, histology (LUAD/LUSC/others), tumor stage (1/2/3/4), vital status (death/live) and follow-up time. 3. Summary statistics of association: SNP-based results with having been adjusted with appropriate variables.    1. For survival: log(HR), standard error of log(HR), 95% CI of log(HR), z value, P value. | |
| **VII. Time line** | |
| Once we receive the datasets from PIs, we need 3 months to do the association analysis, and an additional 3 months to write the manuscript. | |
| **VIII. Other remarks** (e.g. publication plan, etc) | |
| We will plan a manuscript by end of Spring, 2020. | |