Contribution of blood DNA methylation to the association between smoking and lung cancer

Arce Domingo-Relloso,[1] Roby Joehanes,[2] Zulema Rodriguez-Hernandez,[3] Karin Haack,[4] M. Daniele Fallin,[5] Jason G. Umans,[6] Lyle G. Best,[7] Tianxiao Huan,[8] Chunyu Liu,[9] Jiantao Ma,[10] Jose D. Bermudez,[11] Shelley A. Cole,[12] Dorothy A. Rhoades,[13] Daniel Levy,[14] Ana Navas-Acien,[15] and Maria Tellez-Plaza[16]

[1] ad3531@cumc.columbia.edu. Department of Biostatistics, Columbia University, USA.

[2] roby.joehanes@nih.gov. National Heart, Lung, and Blood Institute, USA.

[3] zulema.rodriguez@isciii.es. National Center for Epidemiology, Spain.

[4] khaack@txbiomed.org. Population Health Program, Texas Biomedical Research Institute, TX, USA.

[5] dfallin@jhu.edu. University of Emory, GA, USA.

[6] jgu@georgetown.edu. MedStar Health Research Institute, Washington DC, USA.

[7] lbest@restel.com. Missouri Breaks Industries and Research Inc, SD, USA.

[8] Tianxiao.Huan@umassmed.edu. University of Massachusetts Medical School, MA, USA.

[9] liuc@bu.edu. Boston University School of Public Health, MA, USA.

[10] jiantao.ma@tufts.edu. Tufts University Friedman School of Nutrition Science and Policy, MA

[11] Jose.D.Bermudez@uv.es. Department of Statistics, University of Valencia, Spain.

[12] scole@txbiomed.org. Population Health Program, Texas Biomedical Research Institute, TX, USA.

[13] Dorothy-Rhoades@ouhsc.edu. Stephenson Cancer Center, USA.

[14] levyd@nhlbi.nih.gov. National Heart, Lung, and Blood Institute, USA.

[15] an2737@cumc.columbia.edu. Department of Environmental Health Sciences, Columbia University.

[16] m.tellez@isciii.es. National Center for Epidemiology, Spain.

Altered DNA methylation (DNAm) might be a biological intermediary in the pathway from smoking to cancer. In this study, we investigated the contribution of differential blood DNAm to explain the association between smoking and lung cancer incidence. Blood DNAm was measured in 2321 Strong Heart Study (SHS) participants. Incident lung cancer was assessed as time to event diagnoses. We conducted mediation analysis, including validation using DNAm and paired gene expression data from the Framingham Heart Study (FHS) and bioinformatics analyses to assess the biological plausibility of the findings. We used a novel multiple mediator model (*multimediate*), which is able to assess multiple correlated mediators in a survival outcome setting. In the SHS, current versus never smoking and pack-years single-mediator models showed, respectively, 29 and 21 differentially methylated positions (DMPs) for lung cancer (14 of 20 available, and five of 14 available, replicated, respectively, in FHS) with statistically significant mediated effects. In FHS, replicated DMPs showed gene expression downregulation largely in trans, and were related to biological pathways in cancer. The multimediator model identified that DMPs annotated to the genes *AHRR* and *IER3* jointly explained a substantial proportion of lung cancer. Thus, the association of smoking with lung cancer was partly explained by differences in baseline blood DNAm at few relevant sites. Experimental studies are needed to confirm the biological role of identified eQTMs and to evaluate potential implications for the prevention and control of lung cancer. Our multimediate R package is available in the following Github repository: https://github.com/AllanJe/multimediate.

Keywords: Smoking, DNA methylation, causal mediation analysis.