






Identifying Individual Medications Affecting Pulmonary Outcomes When Multiple Medications are Present

Yisha Li ¹, Ran Dai², Yeongjin Gwon ², Stephen I Rennard³, Barry J Make ⁴, Dinah Foer⁵, Matthew J Strand⁴, Erin Austin⁶, Kendra A Young ¹, John E Hokanson¹, Katherine A Pratte⁷, Rebecca Conway¹, Gregory L Kinney ¹

On behalf of COPDGene investigators

¹Department of Epidemiology, University of Colorado Anschutz Medical Campus, Aurora, CO, USA; ²Department of Biostatistics, School of Public Health, University of Nebraska Medical Center, Omaha, NE, USA; ³Division of Pulmonary, Critical Care and Sleep Medicine, University of Nebraska Medical Center, Omaha, NE, USA; ⁴Department of Medicine, National Jewish Health, Denver, CO, USA; ⁵Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA; ⁶Mathematical and Statistical Sciences, University of Colorado Denver, Denver, CO, USA; ⁷Department of Biostatistics, National Jewish Health, Denver, CO, USA

Correspondence: Gregory L Kinney, Department of Epidemiology, University of Colorado Anschutz Medical Campus, Aurora, CO, USA, Tel +1 303-724-4437, Email GREG.KINNEY@CUANSCHUTZ.EDU

Introduction

Chronic obstructive pulmonary disease (COPD) is a major cause of death globally and is mostly attributable to smoking cigarettes.¹ COPD frequently coexists with other chronic conditions including cardiovascular disease (CVD), metabolic syndrome including type II diabetes, anxiety/depression, etc., known as comorbidities. Previous literature shows patients with COPD and comorbid conditions are associated with increased hospitalization rate, COPD exacerbations, mortality rate and poor quality of life.¹ Patients with COPD may be taking multiple medications in addition to COPD-directed treatment. Therefore, it is important to understand how concurrent medication use beyond COPD treatment might influence pulmonary outcomes. To our knowledge, no previous study has done analysis screening large numbers of concurrent medication use and change of pulmonary outcomes. Our study aims to use a novel, statistical screening approach to identify associations between concurrent non-COPD medication use and longitudinal pulmonary outcomes. This approach could be helpful to identify statistically robust associations between medication classes and change of pulmonary function in patients with or at risk for COPD. From a clinical view, our finding will select concurrent medications other than COPD treatments that might have additional impact on lung function and pathologic tissue change.

There are some challenges in identifying potential associations between concurrent medication use and pulmonary outcomes in populations with or at risk of COPD. First, the candidate feature collection is large. Second, the signals are likely to be rare and weak because the medications are not primarily treating lung diseases. Powerful feature selection procedures for high-dimensional signal identification with reproducibility guarantees are highly desirable. Recently developed methodology in controlling the false discovery rate (FDR) such as the Knockoff methods address these concerns.^{2,3} The FDR is defined as the expected ratio between the number of false discoveries and total discoveries, which conceptualizes the rate of type I errors.² FDR control methods offer statistical guarantees for reproducibility to identify signals across multiple experiments.³ The Knockoff methods construct a knockoff copy of the feature matrix to mimic its correlation structure and serve as a negative control, allowing identification of the true signals with finite sample statistical guarantee of FDR control and higher power compared with other methods (for example, Benjamini-Hochberg method).² The Knockoff methods can work with high-dimensional feature settings which is a desirable method to find the true signals among a large number of variables.² Knockoff methods can pair with many variable selection procedures, including some black-box machine learning algorithms. Least absolute shrinkage and

selection operator (LASSO) regression is a least square method with a penalty on the L1 norm of the coefficient vector that forces regression coefficients to zero unless there is strong evidence of an association.² Compared to some other penalized regressions such as L0 penalized regression, LASSO is computationally efficient.

In our study, we used the Knockoff filter paired with LASSO regression to identify medications that are associated with pulmonary outcomes among smokers with or at risk for COPD with FDR control guarantee. This approach will efficiently select the true signals among large numbers of medication candidates that are associated with change of pulmonary outcomes.

Methods

Data were collected from the COPDGene, a cohort of current or ex-smokers with ≥ 10 pack-years smoking history at baseline (phase 1, P1). The features of the study cohort and protocol have been previously described.⁴ Participants from the COPDGene cohort who completed P1, and 5-year follow-up Phase 2 (P2) from 2007 through 2017 with complete demographics, clinical information, P1 and P2 medication records were included in the study. Participants' medications were collected through self-report and confirmed by bringing their bottles of medications prescribed by their physicians to study visits. Medication classification was based on expert opinion.⁵ We described numbers of people who used more than 2 classes of medications at baseline to indicate the multiple medications environment. Our final cohort included 3954 participants who reported medication use at both P1 and P2 and had complete COPDGene visit data.

Statistical Analysis

We aimed to identify medications associated with pulmonary outcomes using model selection for the difference between P1 and P2 of four outcomes associated with COPD including lung function measures and pathologic lung tissue changes. Lung function change was measured using absolute forced expired volume in one second in millimeters (FEV1), FEV1% predicted (FEV1_pp) using the Hankinson method, forced vital capacity in liters (FVC). Pathological lung change was measured using percent emphysema (Emphysema). LASSO regression was used to perform a variable selection from 42 medication classes at P1 and 38 medication classes at P2 to identify classes associated with pulmonary outcomes. To provide reproducibility, we used the Knockoff method to choose the LASSO tuning parameter to control the FDR level at 0.2. The FDR controlled at 0.2 has been commonly used in previous studies.⁶ We conducted a stability selection procedure, where we constructed 100 different knockoff matrices to perform feature selection 100 times among all medication classes. All 100 selections are with FDR guarantee. However, selection results are slightly different because of the different knockoff constructions. We further summarized the selection frequency of each medication class from the stability selection as a reliability measure of the medication class. The higher frequency a medication class has been selected, the more likely it is a true signal. We considered frequency of stability selection >0 as having a signal, and frequency of stability selection >0.5 as a strong association in our results. Medication classes did not include medications primarily treating COPD such as bronchodilators, glucocorticoids, and antibiotics. Medication classes are dichotomous variables reported as yes/no when medication is present. We included baseline information of gender, race, age, height, FEV1/FVC, FEV1, smoking status, and smoking pack-years in the model.

We used the R package glmnet for LASSO selection. We used the software Knockoffs for the Knockoff method with FDR control guarantees.²

Results

The Knockoff method paired with LASSO identified 13 out of 42 classes of medications in P1 and 13 out of 38 classes of medications in P2 as associated with pulmonary outcomes. Forty-five percent of the participants in our cohort used at least 2 classes of medications at P1 (beyond COPD-directed treatments). Selection frequency of all classes of medications are shown in [Appendix Table 1](#). All selected medications and outcomes are shown in [Appendix Table 2](#). [Figure 1](#) shows that there were 7 classes of medications strongly associated with the four pulmonary outcomes based on frequency of stability selection >0.5 for at least one outcome. There were 4 classes of medications both selected in P1 and P2, which include dopamine reuptake inhibitor, resins, diuretic, and proton pump inhibitor ([Appendix Table 2](#)). P2 tricyclic antidepressant (TCA), P2 diuretic, P2 central agonist and P1 biguanides were strongly associated with the change of FVC from P1 to P2 ([Figure 1](#)). P2 TCA, P2 diuretic, P1 selective cholesterol absorption inhibitors, and P1 biguanides were

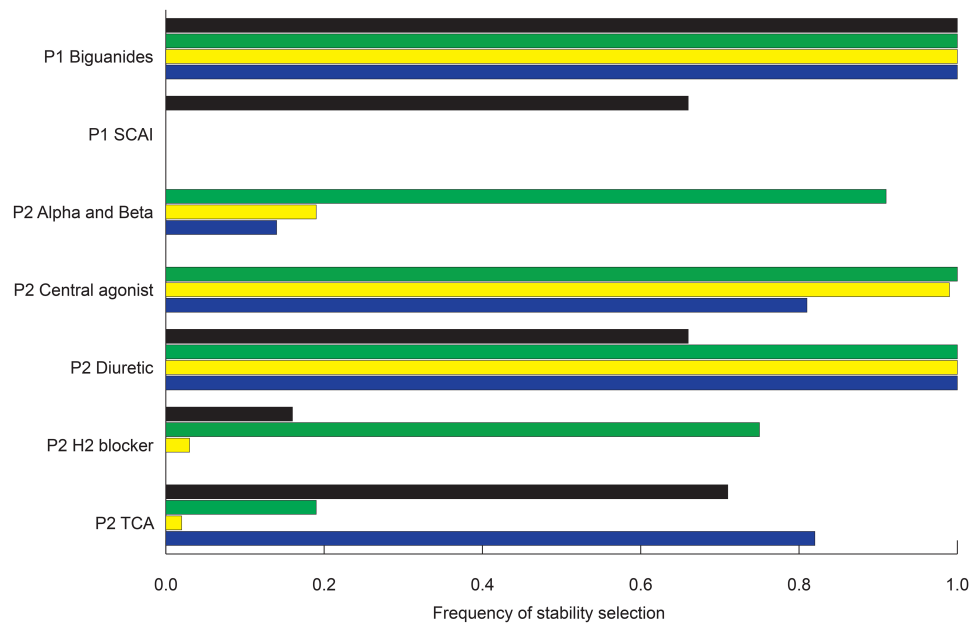


Figure 1 All selected medications meeting the cutoff criteria for one or more pulmonary outcomes. Figure 1 shows all 7 concurrent medications strongly associated (frequency of selection stability >0.5 for at least one outcome) with Emphysema (black bar), FEV1 (green bar), FEV1_pp (yellow bar) and FVC (blue bar) progression. The x-axis of Figure 1 is frequency of stability selection experiment using the Knockoff variable selection procedure. Higher frequency from the stability selection means a particular variable is more likely to be a true signal.

Abbreviations: LASSO, least absolute shrinkage and selection operator; P1, phase 1 (baseline); P2, phase 2 (5-year follow-up); Emphysema, percent of emphysema; FEV1, absolute forced expired volume in one second in millimeters; FEV1_pp, forced expiratory volume in one second in percent predicted using the Hankinson method; FVC, forced vital capacity in liters; Alpha and Beta, combined alpha and beta blockers; TCA, tricyclic and tetracyclic antidepressants; SCAI, selective cholesterol absorption inhibitors.

strongly associated with the change of Emphysema from P1 to P2 (Figure 1). P2 H2 blocker, P2 diuretic, and P1 biguanides were strongly associated with the change of FEV1 from P1 to P2 (Figure 1). P2 diuretic, P2 central agonist, and P1 biguanides were strongly associated with the change of FEV1_pp from P1 to P2 (Figure 1). Biguanides at P1 was the class of medications most strongly associated with all the selected outcomes from P1 to P2 with a selection of 1.

Discussion

Our study used a Knockoff method to select 26 classes of medications associated with longitudinal pulmonary outcomes from P1 to P2 in the COPDGene cohort. To our knowledge, this is the first study that screened all available classes of medications other than COPD treatments associated with COPD progression in a cohort of smokers. Our findings identified classes of medications that are potentially associated with COPD progression among smokers. Our approach was novel in using the well-established Knockoff method to control the FDR paired with LASSO regression to select signals from a big number of medication candidates. The FDR controlling procedure provides a reproducibility guarantee and efficiency when selecting large numbers of medication class candidates.³ This approach can be generalized to other studies conducting variable selections with similar conditions. This finding will be useful for future studies understanding which concurrent medications other than COPD treatments may benefit COPD progression. From a clinical perspective, our work underscores the importance of considering concurrent medications and comorbidities as part of our patients' global COPD care. Future identification of statistically robust associations between medication classes and change of pulmonary outcomes in patients with or at risk for COPD may inform treatment selection and clinical guidelines for COPD patients.

We found biguanides at P1 was the strongest signal associated with all selected pulmonary outcomes from P1 to P2 in COPDGene cohort. Notably, the only biguanides medication in the COPDGene is metformin. A potential mechanism between metformin and pulmonary function was published previously that metformin may act against cigarette smoking-induced pulmonary pathologies as an AMP-activated protein kinase activator.⁷ Our findings suggesting a longitudinal association between metformin and potential change of pulmonary function in the COPDGene should be investigated in

additional prospective studies. Our methods additionally reported other selected classes of medications associated with lung function and pathologic lung tissue change such as TCA, central agonist, diuretic, cholesterol absorption inhibitors, etc. (Figure 1). These selected medications and pulmonary outcomes should be investigated.

Our study had several limitations. First, our study only examined four pulmonary outcomes of interest. However, our Knockoff method can be applied to other outcomes in the future, supporting its utility in research. Second, our findings were based on a population of smokers from COPDGene. While we did not apply it to other cohort-based datasets, the FDR control provides reproducibility guarantee, supporting generalizability, and we would therefore expect to find similar signals in other cohorts with similar characteristics.³ Our study is a preliminary screening analysis and was specifically conducted across the full spectrum lung function and pathologic tissue change. COPD status is based on a cutoff value FEV1/FVC <0.7, and treatment is often contingent on diagnosis. Future analysis will investigate the complex relationship between COPD diagnosis, COPD treatment and concurrent medication use. Our novel method identified medication classes associated with COPD progression, providing a basis for future mechanistic and translational studies to investigate how concurrent medication use impacts COPD, and a model to study associations between medication exposure and more broadly diseases. Third, both P1 and P2 medication use represent people who started the medication before the assessment, however we do not have information about when people started their medications at each phase, thus our results only indicated the association between medication use and pulmonary outcomes rather than causality.

Conclusions

We used a novel and efficient statistical approach to identify a total of 26 classes of medications that are associated with lung function and pathologic lung tissue change. Our findings are useful in identifying concurrent medication use related to changes in COPD. Our findings will also narrow down the medication candidates and therapeutic targets for future observational studies or clinical trials to detect new treatments or potential preventive approaches of COPD.

Abbreviations

COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; Emphysema, percent emphysema; FDR, false discovery rate; FEV1, absolute forced expired volume in one second in millimeters; FEV1_pp, FEV1% predicted using the Hankinson method; FVC, forced vital capacity in liters; LASSO, least absolute shrinkage and selection operator; P1, Phase 1; P2, Phase 2; TCA, tricyclic and tetracyclic antidepressants.

Data Sharing Statement

Data supporting this manuscript are available to all COPDGene investigators. Data are available upon request through <https://dccweb.njhealth.org/sec/COPDGene/MainPage.cfm>.

Ethics Approval and Consent Form

All study participants were informed about the purpose of the study. Confirmation of consent forms were received from study participants. Protocols for all phases (HS-2778) and data evaluation in the COPDGene cohort followed the principles of the Declaration of Helsinki and was approved by the COPDGene committee.

Consent for Publication

We confirmed that details of words, tables, figures included in the manuscript can be published.

Acknowledgment

We thank COPDGene investigators who helped acquiring the data for this study. COPDGene is supported by Award Number U01 HL089897 and Award Number U01 HL089856 from the National Heart, Lung, and Blood Institute. COPDGene is also supported by the COPD Foundation through contributions made to an Industry Advisory Board that has included AstraZeneca, Bayer Pharmaceuticals, Boehringer-Ingelheim, Genentech, GlaxoSmithKline, Novartis, Pfizer, and Sunovion. We thank the Colorado School of Public Health Laboratory for Analytical and Computational Epidemiology (LACE) for their support.

Author Contributions

All authors made a significant contribution to the work reported. All authors took part in drafting, revising or critically reviewing the article, whether that is in the conception of study design, execution, acquisition of data, analysis, and interpretation of the results. All authors gave final approval of the version to be published and have agreed on the journal to which the article has been submitted and agreed to be accountable for all aspects of the work.

Funding

The work was supported by award number U01 HL089897 and award number U01 HL089856 from the National Heart, Lung, and Blood Institute.

Disclosure

In the last three years, Dr. Stephen I. Rennard has received fees for consulting from Alpha 1 Foundation, Berfenbio, GlaxoSmithKline, Verona, Boehringer Ingelheim, Novo Ventures and Sanofi. He is also the Medical Director for the Alpha 1 Foundation Therapeutic Development Network. From 2015 to 2019, he was an employee of AstraZeneca and received shares as part of his compensation. Over the last three years Dr. Barry Make reports grants from NHLBI, American Lung Association, Department of Defense, Pearl Research, Circassia, GlaxoSmithKline, and AstraZeneca; advisory board fees from GlaxoSmithKline, AstraZeneca, Boehringer Ingelheim, Mt. Sinai, Web MD, Novartis, American College of Chest Physicians, Projects in Knowledge, Novartis, personal fees from American College of Chest Physicians, personal fees from Projects in Knowledge, Mylan, Eastern Pulmonary Society, Wolters Kluwer Health (Up-To-Date), Optimum Patient Care Global Limited, Integritas Communications, Quintiles, University of Wisconsin, Third Pole, and Phillips; consulting fees from AstraZeneca; Data Safety and Monitoring Board fees from NHLBI, Takeda and Spiration. Dr Kendra A Young reports grants from NIH, during the conduct of the study. The authors report no other conflicts of interest in this work.

References

1. Negewo NA, Gibson PG, McDonald VM. COPD and its comorbidities: impact, measurement and mechanisms. *Respirology*. 2015;20(8):1160–1171. doi:10.1111/resp.12642
2. Barber RF, Candès EJ. Controlling the false discovery rate via knockoffs. *Ann Stat*. 2015;43(5):2055–2085, 2031. doi:10.1214/15-AOS1337
3. Dai Ran ZC. Multiple testing for composite null with FDR control guarantee. *arXiv preprint arXiv*. 2021. doi:10.48550/arXiv.2106.12719
4. Regan EA, Hokanson JE, Murphy JR, et al. Genetic epidemiology of COPD (COPDGene) study design. *Copd*. 2010;7(1):32–43. doi:10.3109/15412550903499522
5. Li Y, Ragland M, Austin E, et al. Co-morbidity patterns identified using latent class analysis of medications predict all-cause mortality independent of other known risk factors: the COPDGene[®] Study. *Clin Epidemiol*. 2020;12:1171–1181. doi:10.2147/CLEP.S279075
6. Shen A, Fu H, He K, Jiang H. False discovery rate control in cancer biomarker selection using knockoffs. *Cancers*. 2019;11(6):744. doi:10.3390/cancers11060744
7. Polverino F, Wu TD, Rojas-Quintero J, et al. Metformin: experimental and clinical evidence for a potential role in emphysema treatment. *Am J Respir Crit Care Med*. 2021;204(6):651–666. doi:10.1164/rccm.202012-4510OC

Clinical Epidemiology

Dovepress

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification, systematic reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, and evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>