Category: Computational Biology & Bioinformatics

Research Plan **Student Name:** Catherine Kim

Project Title: Co-administration of Atorvastatin Blocks CYP3A4: Exacerbated Risk of Interstitial Lung Disease

A. RATIONALE

Drug-drug interactions involving statins often cause severe adverse drug events (ADE) [1-3]. Statin-drug interactions are of particular importance because over 221 million statin prescriptions are dispensed to over 40 million individuals nationwide and these statins are also frequently co-administered with other drugs, which may ultimately result in harmful drug-drug interactions in patients [4-6]. Current literature reveals contradicting results about the association between statins and interstitial lung disease (ILD) as an ADE. While a systematic review of published case reports and the FDA adverse events reporting system (FAERS) supports the association between statin administration and ILD, past cohort studies of FAERS does not support this finding [7-10]. Along with these contradictions in science, the lack of knowledge on the mechanism of statin-induced ILD also highly restricts current understanding of interstitial lung disease as an ADE of statin administration [11-13]. A major flaw in these studies regarding the association between statin administration and interstitial lung disease is the lack of consideration of Cytochrome P450 (CYP) enzymes in statin-drug interactions leading to ILD [14]. Although it is supported that the co-administration of statins and CYP drug inhibitors increases the likelihood of ADEs in general, this trend remains unclear in the relationship among statins, ILD, and CYP enzymes [15-17]. Thus, the statistical and molecular connection between the co-administration of statins and CYP-inhibiting drugs and ILD must be studied further to clarify current contradictions in science and pave the way for safer methods of statin prescription to ultimately prevent ILD as an ADE in patients.

B. RESEARCH QUESTION(S), HYPOTHESIS(ES), ENGINEERING GOAL(S), EXPECTED OUTCOMES

I. Research Questions

a. How are drug-drug interactions involving the co-administration of statins with other drugs associated with ILD through CYP enzymatic action?

b. How do other compounds with unknown inhibition to CYP enzymes affect risk of ILD when co-administered with statins?
c. What biological mechanisms involving the off-target binding of statins are related to statin-induced ILD?

II. Hypotheses

a. If the difference of means of the proportional reporting ratios (PRRs) of statin and drug inhibitor combinations with common CYP interaction and statin and drug inhibitor combinations with no common CYP interaction result in significant p values for only combinations that target CYP3A4, then statistical associations between the co-administration of statins and CYP3A4-inhibiting drugs and ILD exist.
b. If the natural compounds bind to the same CYP active site as statins and CYP-inhibiting drugs, then they pose an increased likelihood of causing ILD when co-administered with statins.
c. If statins are not sufficiently metabolized by CYP enzymes due to enzyme competition, the statins will bind to an off-target to ultimately cause ILD.

III. Engineering Goals

The objective of this study will be to elucidate the prognosis of ILD with the co-administration of statin and drug inhibitor combinations based upon interactions with CYP enzymes.

IV. Expected Outcomes

a. A significant association will exist between the co-administration of statins and CYP3A4-inhibiting drugs and ILD, but not between the co-administration of statins and CYP-inhibiting drugs and ILD collectively.
b. Natural compounds with a chemical similarity to CYP-inhibiting drugs shown to possess statistical associations to statin-induced ILD will bind to the same CYP active site and thus, also have an increased likelihood of causing ILD when taken with statins.

c. A biological network among significant off-target proteins of statins highly expressed in the lung, proteins in the surfactant metabolism, and interacting proteins will be created around the off-target binding of statins to the off-target, which will then disrupt surfactant metabolism and cause ILD.

C. PROCEDURES, RISK AND SAFETY, DATA ANALYSIS

I. Procedures

*Note: all computational algorithms will be run on a 2017 27-inch iMac on Jupyter Notebook 6.0.2 on Anaconda Navigator 1.9.7 [18].

1. Statistical associations to ILD

1.1 Download FDA Adverse Events Reporting System (FAERS) Dataset

a. The publicly available FAERS dataset will be downloaded from the FDA website from 2004 to the first half of 2019 [19].

1.2 PRR Calculations of Individual Statins

a. For each individual statin, for loops will be coded to run through FAERS data and collect the following: the number of reports where the patient was a) administered the particular statin and acquired ILD, b) administered any other drug and acquired ILD, c) administered the particular statin and did not acquire ILD, and d) not administered the particular statin and acquired ILD using Glob 10.7, NumPy 1.17, SciPy 1.3.3 [18].

b. PRRs will be calculated using the formula $\frac{a/(a+c)}{b/(b+d)}$ with Python 3.6.5 [18].

1.3 Extraction of Data from FAERS

a. Each individual statin will be annotated in terms of metabolizing CYP enzymes using DrugBank 5.1.4. [20].

b. CYP enzymes will be annotated in terms of drug inhibitors using DrugBank 5.1.4. [20].

c. All annotations will be stored in a matrix in a comma-separated values (CSV) file on Microsoft Excel Office 365.

d. Two separate sets of statin and drug combinations will be created using the annotations: statin and drug inhibitor pairs that do not share a common CYP interaction and statin and drug inhibitor paris that share a common CYP interaction.

1.4 PRRs of Statin and Drug Inhibitor Combinations

a. For each statin and drug inhibitor combination, for loops will be coded to run through FAERS data and collect the following: the number of reports where the patient was a) administered the particular statin and drug inhibitor combination and acquired ILD, b) administered any other drug combination and acquired ILD, c) administered the particular statin and drug inhibitor combination and did not acquire ILD, and d) not administered the particular statin and drug inhibitor combination and acquired ILD using Glob 10.7, NumPy 1.17, SciPy 1.3.3 [18].

b. PRRs will be calculated using the formula $\frac{a/(a+c)}{b/(b+d)}$ with Python 3.6.5 [18].

c. PRRs will be grouped based upon common CYP interaction between the statin and CYP inhibitor.

d. Distributions of PRRs for each CYP interaction will be graphed in a boxplot using Microsoft Excel Office 365.

e. PRRs will be converted into log(PRRs) in base 10.

f. Distributions of log(PRRs) for each set of statin and drug inhibitor combinations will be graphed with a histogram and stats.gamma fit on density plots using Seaborn 0.9.0, Pandas 0.25.3, SciPy 1.3.3, and Matplotlib 3.1.1 [18].

g. Top 5 drugs with highest PRRs when co-administered with a statin will be recorded with their corresponding PRR.

2. *Cheminformatics*

2.1 Identification of Natural Products from ChEMBL

a. The drug inhibitors with PRRs greater than 1 for ILD when co-administered with statins will be recorded.

b. All natural products with a chemical similarity to the recorded drug inhibitors from the ZINC library will be recorded [21].

2.2 Molecular Docking with SwissDock

a. CYP enzyme structure will be downloaded from RCSB PDB [22].
b. The HEM ligand will be manually removed from the CYP enzyme with UCSF Chimera 1.5 [23].
c. All recorded drug inhibitors and all recorded natural products will be prepared for docking using the ZINC library [24].
d. Each individual statin, all recorded drug inhibitors, and all recorded natural products will be molecularly docked to the CYP enzyme using SwissDock [25].
e. The molecular docking will be visualized with UCSF Chimera 1.5 [23].
f. Natural products that bind to the same active site as the statins and the chemically similar drug inhibitor will be recorded.

3. *Pathway Analysis via Biological Network*

3.1 Identification of Key Protein-Protein Interactions with ChEMBL, Gene 234.0, Reactome 70, BioGRID 3.5, Protemoics DB 3.0, CORUM 3.0, STRING 11.0 [21-22, 24, 26-27]

a. All off-targets of statins from ChEMBL will be recorded [21].
b. The gene expression of each off-target in the lung will be recorded from the NCBI Gene database [26].
c. Off-targets not expressed in the lung will be excluded from the dataset.
d. Proteins in the surfactant metabolism will be recorded from Reactome 70 [27].
e. Interacting proteins between the significant off-targets of statins and proteins in the surfactant metabolism will be recorded from BioGRID 3.5, Proteomics DB 3.0, CORUM 3.0, and STRING 11.0 [21-22, 24, 26-27].
f. All protein/gene interactions will be visualized using Cytoscape 3.7.0 [28].

3.2 Fine-tuning of Biological Network

a. The single compounds reported in FAERS with PRRs greater than 1 for the ADE of small-cell lung carcinoma [19].
b. If the number of PRRs greater than 1 for both ILD and small-cell lung carcinoma is significant, oncogenes and tumor proteins will be kept in the biological network. If not, oncogenes and tumor proteins will potentially be removed from the biological network.

1. **Human participants research:**

   Not applicable to this study.
2. **Vertebrate animal research:**

   Not applicable to this study.
3. **Potentially hazardous biological agents research:**

   Not applicable to this study.
4. **Hazardous chemicals, activities & devices:**

   Not applicable to this study.

II. Risk and Safety

Because all experimentation will be done computationally, this study will pose no risk.

III. Data Analysis

For significance tests throughout the study, the mean and standard deviation will be calculated for both sets of statin and drug inhibitor combinations and 2-sample z tests at an $\alpha = 0.05$ level will be performed.

Addendums

1. ***Statistical associations to ILD (cont.)***

1.4 PRRs of Statin and Drug Inhibitor Combinations (cont.)

h. Graph the PRRs for each statin and CYP drug inhibitor pair against the PRR each individual CYP drug inhibitor in a scatter plot using Microsoft Excel Office 365.

4. ***Proof of Concept***

4.1 Statistical Associations to Myopathy

a. Using previously extracted files and values from the FAERS database, calculate PRRs using the formula $\frac{a/(a+c)}{b/(b+d)}$ with Python 3.6.5 for the ADE of myopathy [18-19]

b. Group PRRs based upon common CYP interaction -- CYP2D6, CYP3A4, and CYP3A5 -- between the statin and CYP inhibitor based on literature review.

c. Convert PRRs into log(PRR)s in base 10.

d. Graph distributions of log(PRR)s for each set of statin and drug inhibitor combinations with a histogram and a stats.gamma fit on density plots using Seaborn 0.9.0, Pandas 0.25.3, SciPy 1.3.3, and Matplotlib 3.1.1 [18]

4.2 Literature Support

a. Record the top 6 drugs co-administered with statins with the highest PRRs with their corresponding PRR.

b. Complete a literature review to support that the co-administration between the statin and recorded drug is related to/increases the risk of myopathy.

4.3 Pathway Analysis

a. Record all targets of statins from ChEMBL [21]

b. Determine tissue-specific protein expression from NCBI Gene 234.0 data [26]

c. Determine and record all gene/protein pathway and interactions between proteins using Reactome 70, BioGRID 3.5, Proteomics DB 3.0, STRING 11.0, and CORUM 3.0 [21-22, 24, 26-27]

d. Visualize all protein-protein interactions using Cytoscape 3.7.0. [28]
e. Complete a literature review to support the role of the recorded protein-protein interactions in statin-induced myopathy.

D. BIBLIOGRAPHY

References

[1] C. Zhan, E. Roughead, L. Liu, N. Pratt, and J. Li, "A data-driven method to detect adverse drug events from prescription data," Journal of Biomedical Informatics, vol. 85, pp. 10–20, Sept. 2018.

[2] M. Kovačević, S. V. Kovačević, S. Radovanović, P. Stevanović, and B. Miljković, "Adverse drug reactions caused by drug–drug interactions in cardiovascular disease patients: introduction of a simple prediction tool using electronic screening database items," Current Medical Research and Opinion, pp. 1–11, Aug. 2019.

[3] S. H. D. Hernández and I. Cruz-Gonzalez, "Incidence and preventability of medication errors and ADEs in ambulatory care older patients," The Consultant Pharmacist, vol. 33, no. 8, pp. 454–466, Jan. 2018.

[4] W. S. Weintraub, "Perspective on Trends in Statin Use," JAMA Cardiology, vol. 2, no. 1, pp. 11-12, Jan. 2018.

[5] E. Gaudette, "Did Statins Reduce the Health and Health Care Costs of Obesity?," Pharmacoeconomics, vol. 33, no. 7, pp. 723-734, Jan. 2015.

[6] P. A. Wijnen, O. Bekers, and M. Drent, "Relationship between drug-induced interstitial lung diseases and cytochrome P450 polymorphisms," Current Opinion in Pulmonary Medicine, vol. 16, no. 5, pp. 496–502, Sep. 2010.

[7] M. Schwaiblmair, "Drug induced interstitial lung disease," The Open Respiratory Medicine Journal, vol. 6, no. 1, pp. 63–74, 2012.

[8] A. L. Olson, A. H. Gifford, N. Inase, E. R. F. Pérez, and T. Suda, "The epidemiology of idiopathic pulmonary fibrosis and interstitial lung diseases at risk of a progressive-fibrosing phenotype," European Respiratory Review, vol. 27, no. 150, p. 180077, 2018.

[9] "Interstitial lung disease," Genetic and Rare Diseases Information Center. [Online]. Available: https://rarediseases.info.nih.gov/diseases/13336/interstitial-lung-disease. [Accessed: 01-Oct-2019].

[10] O. Kalchiem-Dekel, J. Galvin, A. Burke, S. Atamas, and N. Todd, "Interstitial lung disease and pulmonary fibrosis: A practical approach for general medicine physicians with focus on the medical history," Journal of Clinical Medicine, vol. 7, no. 12, p. 476, Nov. 2018.

[11] N. Saad, P. Camus, S. Suissa, and P. Ernst, "Statins and the risk of interstitial lung disease: a cohort study," Thorax, vol. 68, no. 4, pp. 361–364, Jan. 2013.

[12] S. Ramkumar, A. Raghunath, and S. Raghunath, "Statin therapy: Review of safety and potential side effects," Acta Cardiologica Sinica, vol. 32, no. 6, pp. 631-639, Nov. 2016

[13] A.-B. Fernandez, R.-H. Karas, A.-A. Alsheik, P.-D. Thompson, "Statins and interstitial lung disease: a systematic review of the literature and of food and drug administration adverse event reports," Chest, vol. 134, no. 4, pp. 824-830, Aug. 2008

[14] U. M. Zanger and M. Schwab, "Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation," Pharmacology; Therapeutics, vol. 138, no. 1, pp. 103–141, Apr. 2013.

[15] C. Fong, "Statins in therapy: Cellular transport, side effects, drug-drug interactions and cytotoxicity- the unrecognized role of lactones," Jul. 2016.

[16] M. Zhelyazkova-Savova, S. Gancheva, and V. Sirakova, "Potential statin-drug interactions: prevalence and clinical significance," SpringerPlus, vol. 3, no. 1, Mar. 2014.

[17] N. C. Ward, G. F. Watts, and R. H. Eckel, "Statin toxicity," Circulation Research, vol. 124, no. 2, pp. 328–350, Jan. 2019.

[18] "Anaconda Navigator," *NC State University Libraries*. [Online]. Available: https://www.lib.ncsu.edu/software/anaconda-navigator. [Accessed: 26-Nov-2019].

[19] *openFDA*. [Online]. Available: https://open.fda.gov/data/faers/. [Accessed: 26-Nov-2019].

[20] D. S. Wishart, Y. D. Feunang, A. C. Guo, E. J. Lo, A. Marcu, J. R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, N. Assempour, I. Iynkkaran, Y. Liu, A. Maciejewski, N. Gale, A. Wilson, L. Chin, R. Cummings, D. Le, A. Pon, C. Knox, and M. Wilson, "DrugBank 5.0: a major update to the DrugBank database for 2018," *Nucleic Acids Research*, vol. 46, no. D1, Aug. 2017.

[21] A. Gaulton, A. Hersey, M. Nowotka, A. P. Bento, J. Chambers, D. Mendez, P. Mutowo, F. Atkinson, L. J. Bellis, E. Cibrian-Uhalte, M. Davies, N. Dedman, A. Karlsson, M. P. Magarinos,

J. P. Overington, G. Papadatos, I. Smit, A. R. Leach, "The ChEMBL database in 2017," *Nucleic Acids Research,* vol. 45, no. 1, pp. 945-954, Jan. 2017.

[22] C. Markosian, Lu. D. Costanzo, M. Sekharan, C. Shao, S. K. Burley, C. Zardecki, " Analysis of impact metrics for the Protein Data Bank," *Scientific Data,* Oct. 2018.

[23] G. S. Couch, "Nucleic acid visualization with UCSF Chimera," Nucleic Acids Research, vol. 34, no. 4, 2006.

[24] F. Fu and D. F. Voytas, "Zinc Finger Database (ZiFDB) v2.0: a comprehensive database of C2H2 zinc fingers and engineered zinc finger arrays," *Nucleic Acids Research*, vol. 41, no. D1, 2012.

[25] K. A. Ramsbottom, D. F. Carr, A. R. Jones, and D. J. Rigden, "Critical assessment of approaches for molecular docking to elucidate associations of HLA alleles with adverse drug reactions," *Molecular Immunology*, vol. 101, pp. 488–499, 2018.

[26] T. Hampton, "NIH Funds Creation of Database of Cellular Signatures," *Jama*, vol. 312, no. 16, p. 1625, 2014.

[27] S. Jupe, "Deubiquitination," *Reactome - a curated knowledgebase of biological pathways*, vol. 57, 2016.

[28] J. H. Morris, S. Lotia, A. Wu, N. T. Doncheva, M. Albrecht, A. R. Pico, and T. E. Ferrin, "setsApp for Cytoscape: Set operations for Cytoscape Nodes and Edges," *F1000Research*, vol. 3, p. 149, May 2015.