ARIC Manuscript Proposal # 3629

PC Reviewed: 5/12/20	**Status: _____**	**Priority: 2**
SC Reviewed: _____	**Status: _____**	**Priority: ____**

1.a. Full Title: A Proteomic Approach for Investigating the Pleiotropic Effects of Statins in the Atherosclerosis Risk in Communities (ARIC) Cohort.

b. Abbreviated Title (Length 26 characters): Pleiotropy of Statins

2. Writing Group: Bruno Bohn, Pamela Lutsey, Weihong Tang, Jim Pankow, Faye Norby, Christie Ballantyne, Eric Whitsel, Kunihiro Matsushita, Ryan Demmer, other interested investigators are welcome.

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. __BB___ **[please confirm with your initials electronically or in writing]**

First author: **Bruno Bohn**
Address: 1300 South 2nd Street; Suite 300
Minneapolis, MN, 55454

Phone: 612-355-0547
E-mail: ferre083@umn.edu

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).
Name: Ryan Demmer
Address: 1300 South 2nd Street; Suite 300
Minneapolis, MN, 55454

Phone: 612-626-8597
E-mail: demm0009@umn.edu

3. Timeline:
Analyses to begin immediately; draft expected in approximately 4 months (August).

4. Rationale:
Statins are a major class of drugs that are primarily prescribed to reduce plasma cholesterol levels. However, statins have been shown to have a variety of other effects in various

biological pathways and systems, including modulation of inflammation and inflammatory cell response, endothelial functioning and nitric oxide (NO) synthesis, and atherosclerotic plaque formation and stability [1, 2]. The mechanisms behind the pleiotropic effects of statins are categorized as lipid-dependent or lipid-independent [3, 2], and may vary according to statin type and dosage [2].

Classically, the lipid-dependent pleiotropic effects of statins have been linked to a decrease in cholesterol synthesis or the increase in cholesterol removal from circulation [3, 2, 1]. LDL-C also has a role in increasing inflammation through LDL-mediated activation of transmembrane receptors (ex: Toll Like Receptors, TLRs), which lead to exacerbated inflammatory responses [3]. LDL-C may also lead to the formation of atherosclerotic plaque through increasing platelet activity and driving macrophages to form pro-atherogenic cells [3]. Further, cholesterol crystals prime nod-like receptor protein 3 (NLRP3) inflammasomes for activation of interleukin (IL)-1β, a pro-inflammatory cytokine [3].

Decrease in LDL-C levels has been associated to a restoration in endothelial function and improvement in vasodilation, likely through the negative impact on oxidized forms of LDL-C have eNOS and downregulation of NO synthesis [2]. Increased endothelial NO synthesis has a significant role in various biophysiological processes, including a decreased risk in cardiovascular illnesses, a decrease in inflammatory response, and inhibition of platelet aggregation [3, 2].

Statins primarily act as inhibitors of the enzyme HMG-CoA reductase through competitive binding to its active site [3, 2]. HMG-CoA reductase is an integral component in the isoprenoid pathway (also referred to as the mevalonate pathway), which takes place primarily in hepatic cells [2]. Thus, statins also impact the synthesis of molecules downstream in the isoprenoid pathway through limiting the activity of HMG-CoA reductase [3, 2, 1].

Inhibition of the isoprenoid pathway by statins limits the synthesis of isoprenoids, including the lipids Farnesylpyrophosphate (FPP) and Geranylgeranylpyrophosphate (GGPP) [2]. These lipids are involved in isoprelynation, a post-translational modification of several proteins – including signal transduction molecules, including the small guanosine triphosphate (GTP)-binding proteins [3, 2]. Therefore, statins may be associated with the inhibition of post-translational modification of various proteins, impacting a wide range of biophysiological effects, including pro-atherogenic and pro-inflammatory signaling [3, 2].

The impact of statins in endothelial function are not limited to lowering cholesterol levels and inhibiting isoprenoids synthesis [3, 2]. Statins further lead to eNOS functioning through prolonging its mRNA half-life and by several interaction with proteins involved in eNOS coupling and transcriptional activation and regulation [3, 2]. Statins are associated to increased endothelial health through other mechanisms, including activating expression of tissue-type plasminogen activator (t-PA) and endothelin-1 (ET-1), increasing in endothelial progenitor cell (EPC) circulation, proliferation, and survival [4, 2].

Further, statins have antioxidant effects through inhibiting the production of reactive oxygen species (ROS) through upregulating the expression of ROS scavenging enzymes and

through Rac1-mediated increase in NADH oxidase activity [3]. The statin-mediated reduction of oxidative stress modulates various redox-sensitive transcription pathways, which regulates the expression of pro-inflammatory and pro-atherosclerotic genes [3].

Statins have an impact in increased atherosclerotic plaque stability, reduced plaque size, and modifications in the lipid core [4, 3, 2]. Through the decrease in LDL-C, statins lead to a decrease in macrophage accumulation in atherosclerotic lesions. Additionally, statins have been shown to reduce the number of inflammatory cells in plaques through modulating the production and secretion of various cytokines, chemokines, and monocytes [3, 2]. Statins may also have a role in increasing the number of regulatory T cells , modulating leukocyte trafficking, activating T cells, and inhibiting the interaction between monocytes and the endothelium [3, 2].

The impact of statins on inflammation has also been evidenced through lower plasma levels of the high-sensitive C-reactive protein (hs-CRP), which may occurs through lipid-dependent mechanisms or through the immunomodulatory function of statins [3, 2].

Statin's powerful and pleiotropic effects are clearly established. However, much of the aforementioned literature is based on basic science studies. Though much is known about the pharmacodynamics of statins, it is possible that additional pathways remain to be elucidated. The ARIC SomaScan data provide a unique and unprecedented resource with which to identify new pathways, and bolster evidence for hypothesized and established pathways, through which statins may influence human health and disease.

5. Main Hypothesis/Study Questions:
Statins have pleiotropic effects that will demonstrate significant associations with numerous proteins – both hypothesized and novel – in the plasma proteome. Hypothesized pathways include those related to lipid metabolism, nitric oxide synthesis, inflammation, and oxidative stress.

6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

Design:
Cross-sectional, with data from visit 3 being used in the primary analysis.

We do recognize the shortcomings of having an observational study to observe the effect of a medication. Of note, unmeasured confounding may be introduced due to statin users being inherently different from non-statin users and several forms of pharmacoepidemiologic biases may be introduced. We will utilize pharmacoepidemiologic techniques such as matching and possibly covariate adjustment by propensity scores to minimize these biases.

Visit 3 was selected as the primary data source due to larger sample size and less reported comorbidities overall, including less kidney function impairment. We are, however, cognizant that the number of statin users is somewhat limited (~5%).

Visit 5 could be used as a nuanced replication sample. We considered utilizing visit 5 for the primary analysis, (~50%) statin users, but had concerns about the smaller total sample size, more medication use in general, generalizability given that many with high lipids would have died, and exchangeability of non-statin users.

Inclusion/Exclusion:
Exclusion: Participants who do not have SOMAscan data at visit 3 or with failed QC.

Exposure:
Definition 1: Defined as use of statins of any kind versus none.
Definition 2: Defined by statin type and dosage.
Although dosage was not recorded, it can be estimated following the conversion of the recorded pill size in mg to dose equivalence and a first-order approximation based the assumption that most prescriptions are once daily. This approach was conducted previously, using the WHO ATC defined daily dosages.

Outcome:
Natural log of SomaScan protein levels, after removal of non-human proteins and proteins with unacceptable QC.

Data Analysis:
A propensity score analysis will be utilized to mitigate the issues associated to observational pharmacoepidemiologic studies [6].

A potential approach includes propensity matching statin users to non-users. Propensity matching will be considered of both fixed and time-varying covariates. Fixed covariates may include sex and race/ethnicity, while time-varying covariates may include age, BMI, and clinical characteristics such as LDL-C and comorbidities. Utilizing time-varying covariates will be crucial to minimize pharmacoepidemiologic confounding.

Other potential approaches include using propensity scores as predictors in regression analyses or the use propensity of score weighing, such as the inverse probability of treatment weighted estimation [6]. The latter can incorporate data from previous visits, leveraging the longitudinal nature of the data.

Multivariate Linear Regression will be used to identify differences in protein levels (dependent variable) between those using any statins vs none (primary analysis). Adjustment may be necessary in a propensity weighing analysis, to account for covariates such as sex, race, age, eGFR, and selected medications (ie. cholesterol-lowering medications and antidiabetic medications) or comorbidities. Statistical significance will be determined based on a p-value of 0.05, adjusted for multiple comparisons through FDR.

Regression approaches will primarily use the R package LIMMA (Linear Models for Microarray Data) due to its strengths to detect differential protein levels in microarray studies [8]. Due to the highly parallel nature of genomic data, LIMMA utilizes regression models that are fitted as to account the entire experiment rather than conducting independent tests for each protein. Through

utilizing a Parametric Empirical Bayes (PEB) framework, the residual variance for a specific protein is moderated through that of other proteins, increasing the effective degrees of freedom for estimating variances and differentially "squeezing" variances to improve predicting power and accuracy. Aside from being beneficial to deal with smaller sample sizes, this approach minimizes the number of false positives for genes with small variances, while improving the power to observe differences in protein levels for those with high variances [8].

A number of studies have demonstrated the use of LIMMA to model SomaScan protein data in various settings. Bowedes et. al. (2019) primarily utilized LIMMA to identify novel proteins associated with Sjogren's Syndrome in a case-control study (n = 83) [9]. Similarly, LIMMA was utilized by Kollar et. al. (2018) to model differences in protein levels in face transplant patients with and without tissue rejection [10]. Further, Marion et. al. (2016) utilized LIMMA to characterize the respiratory mucosal proteome in influenza patients, with data derived from SomaScan [11]. Lastly, Cotton and Graumann (2016) have described in more detail its use in conjunction with the R package *readat*, the primary package to access SomaScan data in the platform [12].

A secondary analysis to investigate the effect of dosage by statin type will be conducted utilizing multivariate linear models as described above.

A sensitivity analysis with data from visit 5 (and potentially visit 2, if available) will be performed.

Lastly, the Ingenuity Pathway Analysis (IPA) software will be used to explore the association of statins with biophysiological pathways and processes. The proteins shown to be associated with statin use in linear regression analysis will be used to inform the decisions on which pathways to analyze.

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? ____ Yes __X__ No

b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? ____ Yes ____ No
(The file ICTDER has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? ___ Yes __X__ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the current derived consent file ICTDER05 must be used to exclude those with value RES_DNA = "No use/storage DNA"? ____ Yes ____ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status.

ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/aric/mantrack/maintain/search/dtSearch.html

_____ **Yes** **_X_____** No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?

There are no similar papers.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? __X__ Yes ____ No

11.b. If yes, is the proposal

_X__ A. primarily the result of an ancillary study (list number* _AS2017.27____)

___ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* _____ _____ _____)

*ancillary studies are listed by number https://sites.cscc.unc.edu/aric/approved-ancillary-studies

12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

References

[1] C. R. Sirtori, "The pharmacology of statins," *Pharmacological Research,* pp. 3-11, 2014.

[2] J. K. Liao and U. Laufs, "Pleiotropic effects of statins," *Annual Review of Pharmacology and Toxicology,* pp. 89-118, 2005.

[3] L. Liberale, F. Carbone, F. Montecucco and A. Sahebkar, "Statins reduce vascular inflammation in atherogenesis: a review of underlying molecular mechanisms," *Intl Journal of Biochemistry and Cell Biology,* 2020.

[4] G. Galyfos, A. Sianoy and K. Fillis, "Pleiotropic effects of statins in the perioperative setting," *Ann Card Anaesth,* pp. S43-S48, 2017.

[5] W. A. Ray, "Evaluating Medication Effects Outside of Clinical Trials: New-User Designs," *American Journal of Epidemiology,* vol. 158, no. 9, pp. 915-920, 2003.

[6] M. A. Brookhart, R. Wyss, J. B. Layton and T. Sturmer, "Propensity Score Methods for Confounding Control in Nonexperimental Research," *Circulation Outcomes,* vol. 6, no. 5, pp. 604-611, 2013.

[7] M. C. R. Division of Biomedical Statistics and Informatics, "Biomedical Statistics and Informatics Software," Mayo Clinic, 23 July 2018. [Online]. Available: http://bioinformaticstools.mayo.edu/research/gmatch/. [Accessed 27 April 2020].

[8] M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi and G. K. Smyth, "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acid Research,* p. e47, 2015.

[9] I. L. A. Bodewes, P. J. van der Spek, L. G. Leon, A. J. M. Wijkhuijs, C. G. van Henden-Meeuwsen, L. Tas, M. W. J. Schreurs, P. L. A. van Daele, P. D. Katsikis and M. A. Versnel, "Fatigue in Sjogren's Syntrome: a search for biomarkers and treatment targets," *Frontiers in Immunology,* vol. 10, no. 312, pp. 1-10, 2019.

[10] B. Kollar, A. Shubin, T. J. Borges, S. Tasigiorgos, T. Su Win, C. G. Lian, S. T. Dillon, X. Gu, I. Wyrobnik, G. F. Murphy, B. Pmahac, T. A. LIbermann and L. V. Riella, "Icreased Levels of Circulatin MMP3 Correlate with Severe Rejection in Face Transplantation," *Nature Scientigic Reports,* vol. 8, no. 14915, pp. 1-12, 2018.

[11] T. Marion, H. Elbahesh, P. G. Thomas, J. P. DeVicenzo, R. Webby and K. Schughart, "Respiratory Mucosal Proteome Quantification in Human Influenza Infections," *PLoS One,* vol. 11, no. 4, p. p.e0153674, 2016.

[12] R. J. Cotton and J. Graumann, "readat: an R package for reading and working with SomaLogic ADAT files," *BMC Bioinformatics,* vol. 17, no. 201, pp. 1-5, 2016.