AWARD NUMBER: W81XWH-17-1-0533

TITLE: Metabolomics of Lead Exposure and Its Role in Respiratory Disease

PRINCIPAL INVESTIGATOR: Jessica Lasky-Su

CONTRACTING ORGANIZATION: Brigham and Women's Hospital, Inc.
Boston, MA 02115

REPORT DATE: October 2018

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
October 2018	Annual	15 September 2017-14 September 2018

4. TITLE AND SUBTITLE

Metabolomics of Lead Exposure and Its Role in Respiratory Disease

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-17-1-0533

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jessica Lasky-Su; Rachel Kelly

E-Mail: rejas@channing.harvard.edu; hprke@channing.harvard.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Brigham and Women's Hospital,

181 Longwood Avenue
Boston
MA 02129, USA

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The overarching hypothesis of this proposal was that exposure to lead (Pb) during active military service is related to the observed high prevalence of poor lung health among veterans. We aimed to explore the mechanisms underlying this relationship utilizing metabolomics; i.e. the systematic profiling of small (<10kDa) metabolites in a biological sample, which will allow us to construct a causal pathway demonstrating the mechanistic and biological connections between Pb exposure and lung health. To achieve this aim we identified participants from the ongoing Normative Ageing Study of Veterans. We selected men with detailed histories on their exposure to Pb, with comprehensive data on long term lung health and with blood samples suitable for metabolomics profiling. During this reporting period, we identified 661 plasma samples from 464 veterans, which we shipped to Metabolon Inc. for metabolomic profiling using four LCMS platforms, enabling the broadest coverage of the metabolome possible. We applied QC and data processing pipelines to these data, and initiated the statistical analysis plan outlined in our proposal. To date, we have successfully identified a metabolomic profile associated with Pb exposure and a metabolomic profile associated with poor lung health. Our analyses, encompassing both frequentist and network approaches, suggested that the relationship between Pb and the lung is mediated, in part, by dysregulated Glycine, Serine and Threonine Metabolism; Histidine Metabolism; Leucine, Isoleucine and Valine Metabolism; Phospholipid Metabolism and Sphingolipid metabolism. These findings have been accepted for presentation at the American Thoracic Society Annual conference, and during the next reporting period we aim to publish these initial findings, complete construction of the causal pathway and move toward biomarker development.

15. SUBJECT TERMS

None provided

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	Unclassified	24	**19b. TELEPHONE NUMBER** *(include area code)*
Unclassified	Unclassified	Unclassified			

Table of Contents

Page

1. Introduction

Poor respiratory health relating to environmental exposures during active service represents a significant public health burden for Military service personnel and Veterans. Lead (Pb) is commonly found on military bases in the form of fine particulate matter and is thought to adversely affect pulmonary function for many decades after exposure. However, a complete understanding of the effects of Pb exposure on the respiratory system remains to be fully elucidated, and further investigation is paramount. Metabolomics, the systematic profiling of all the small molecules in a biological system, represents a powerful tool to (i) increase mechanistic understanding of the pathogenesis of Pb exposure on the respiratory system, and (ii) identify biomarkers of 'toxic' levels of exposure. The overarching hypothesis of this proposal is that a metabolomic profile of Pb exposure can be identified and used to understand how Pb exposure and the resulting changes on the metabolome have a downstream impact on respiratory health. Therefore, this projects aims to determine the influence of Pb exposure on respiratory health through the integration of the metabolomic profiles of heavy metal exposure and of respiratory disease by the construction of a biological pathway from exposure to respiratory outcome. This will allow for the identification of novel blood-based biomarkers associated with long-term toxic Pb exposure and respiratory health. These biomarkers will provide mechanistic insights into the pathogenic effects of Pb on the respiratory system and into the disease pathways involved, supporting the development of novel therapeutics. This represents the first study to utilize metabolomics in the exploration of both short-term and long-term Pb exposure and its effect on respiratory disease.

2. Keywords

Respiratory disease; Metabolomics; metabolome; lead (Pb); heavy metals; biomarkers; Metals Toxicology; omics

3. Accomplishments

3.a Major Goals of the Project

The original major goals of this project, as detailed in the **Statement of work (SOW)** in our proposal are show in **Table 1a-c.** These goals are delineated by aim and subdivided by tasks (major and sub), and the extent of their completion and completion dates where relevant, along with accompanying notes is provided. A detailed description of the methodology and findings to date is provided below. The corresponding table or figure describing the results in indicated in square brackets. We note that due to the nature of the analyses, rather than proceeding sequentially as we previously indicated in the **SOW**, it has proven more effective to perform a number of tasks concurrently. For example, the methods and coding to identify the metabolome of respiratory disease and the metabolome of Pb exposure are very similar so we have been undertaking these tasks simultaneously, and consequently, both are at similar stages of completion. One notable issue was a delay in the HPRO approval that has now been resolved (*see section 5a for details*).

Table 1a: SOW and Progress in Accomplishing the Outline Goals (Specific Aim 1)

Specific Aim 1 To identify metabolomic profiles of lead exposure in a population of Veterans	**Timeline**	**% completion**	**Completion Date**	**Comments**
Major Task 1 Metabolomic profiling of 374 NAS individuals	Months			
Subtask 1: Selection and shipment of blood plasma	1	100%	11/16/17	Additional samples; *total 661*, were profiled following price negotiations with Metabolon [**Table 2**]
Subtask 2: Metabolomic profiling	2-3	100%	1/31/18	
Subtask 3: Data processing and Quality control procedures using Metabolon, Inc internal standards	2-3	100%	1/31/18	
Subtask 4: Data Quality Control pipeline at the Channing Laboratory	3	100%	3/31/18	
Milestone Achieved: Metabolomic dataset received ready for analysis	*3*	100%	3/31/18	[**Table 3 & Figure 1**]
HRPO regulatory review	1-3	100%	2/4/18	IRB concluded the research did not constitute human subjects research
Milestone Achieved: HRPO Approval	*3*	100%	2/4/18	
Major Task 2 Identification of the metabolome of Pb exposure				
Subtask 1: Creation of a composite measure of Pb exposure incorporating duration and level, based on pre-existing measures in bone, toenail, blood and urine	1	50%	ongoing	Working with colleagues from NAS to obtain further information on all available measures of lead exposure
Subtask 2: Linear regression models and sensitivity analyses, exploring different confounders, to identify differential metabolites and pathway enrichment analysis to identify the metabolomic pathways mapping to these metabolites associated with exposure	2-3	25%	ongoing	Analyses has been conducted for existing measures of lead exposure including whole blood at the time concurrent to metabolomic profiling and spirometry. Mixed models were used to account for the longitudinal design. Metabolite and pathways of interest were identified [**Table 4; Table 5**]
Subtask 3: Network approaches, including WGCNA to identify metabolomic networks	3-4	25%	ongoing	State of completion is as above [**Figure 2; Figure 3**]. Metabolite networks of interest have been identified
Milestone Achieved: Publication of a manuscript exploring the metabolome of Pb exposure	*5*	10%	ongoing	Systematic literature searches have been conducted, introduction and methods write up are ongoing

Table 1b: SOW and Progress in Accomplishing the Outline Goals (Specific Aim 2)

Specific Aim 2 To identify metabolomic profiles of Pb exposure in a population of Veterans	Timeline	% completion	Completion Date	Comments
Major Task 1 To identify a metabolomic profile of pulmonary function within a population of Veterans				
Subtask 1: Linear regression and network analysis as described in Major task 2 applied to the outcome of pulmonary function, as denoted by FEV_1 and FVC	6-8	75%	ongoing	Analyses nearing completion [**Table 4; Table 5; Figure 2; Figure 3**]
Subtask 2: Development and assessment of a metabolomic score based on the findings from subtask 1 that can be used to discriminate men by their degree of lung function	8-9	20%	ongoing	ROC curve analyses and sensitivity and specificity are used to assess score. To date we have explored generating scores based on the first principal component of the significant metabolites [**Figure 4; Table 6**]. Currently exploring alternative approaches for the development of the score, as well as complementary approaches to assess discriminatory ability.
Milestone Achieved: Publication of a manuscript exploring the metabolome pulmonary function	*9*	10%	ongoing	Systematic literature searches have been conducted, introduction and methods write up are ongoing

Table 1c: SOW and Progress in Accomplishing the Outline Goals (Specific Aim 3)

Specific Aim 3 To determine the influence of Pb exposures on respiratory health through the integration of the metabolomics profiles of Pb exposure and pulmonary function	**Timeline**	**% completion**	**Completion Date**	**Comments**
Major Task 1: Identify metabolites, metabolite networks and metabolomics pathways along the causal pathway from Pb exposure to reduced pulmonary function	Months			
Subtask 1: Identify common differential metabolites and metabolomics pathways for Pb exposure and pulmonary lung function	10	50%	ongoing	Analyses complete for those relating to whole blood Pb exposure. Pending for other Pb measures, final selection of common metabolites will depend on comparison of robustness of the Pb models. [**Figure5; Table 7**]
Subtask 2: Utilize WGCNA preservation analysis to identify the elements of metabolic networks that are consistent for Pb and pulmonary lung function	10-11	0%	pending	
Milestone Achieved: Publication of findings	*12*	10%	ongoing	Abstract accepted to the American Thoracic Society Annual Meeting 2019, to be help in May in Dallas, presenting the results of Major Task1; Subtask1. Abstract additionally submitted to Northeastern University's 2019 Research Innovation and Scholarship Expo (April, Boston)
Major Task 2: Construction of a biologically informative causal pathway				
Subtask 1: Utilize the results from Major Task 1 together with mediation analyses and structural equation modelling to construct a causal pathway from Pb exposure to respiratory outcome	12-14	0%	pending	
Milestone Achieved: Publication of findings	*15*	0%	pending	
Milestone Achieved: Publication discussing the applicability of metabolomics for constructing causal pathways	*16*	0%	pending	
Major Task 2: Biomarker Development				
Subtask 1: Utilizing the results from Aims 1-3 explore the development of biomarkers of exposure, outcome and intermediate biomarkers along the causal pathway	15-18	0%	pending	
Subtask 2: Consider forward translation into a clinical setting, as well as reverse translation into the development of therapeutic targets	17-18	0%	pending	
Subtask 3: Search for suitable replication population(s)	18	0%	pending	
Milestone Achieved: Publication of findings	*18*	0%	pending	
Milestone Achieved: Finalize plan and secure funding for future study to develop findings further	*18*	0%	pending	

3.b. Accomplishments of these goals

There have been no changes to the overall goals of this project as stated in the **SOW** (**Table 1a-c**). The major accomplishments by task are outlined below, method are described and preliminary results and conclusions are presented.

AIM ONE: Major Task one: Metabolomic profiling of NAS individuals

There was an advantageous change in the number of samples we were able to profile following price negotiations with Metabolon Inc., who were sub-contracted to perform the metabolomics profiling analyses (*see section 5.b for further details*). Rather than the 374 samples we stated in our proposal, we were able to profile 661 samples from 464 men, including a number from the same men over a longer period of time (**Table 2**). This substantially increased our overall power, and specifically our power to explore changes in the metabolome over time. Subject selection was carefully performed to ensure we included both a range of lung function and range of Pb exposure, and we prioritized the selection of men with multiple longitudinal blood samples for profiling.

Table 2: Baseline characteristics of 464 Men from the Normative Ageing Study with Metabolomic of Plasma Samples

Characteristic		n=464 Men from the Normative Ageing Study
Age (yrs)	**mean [range]**	75 [57, 97]
BMI	**Underweight *n* (%)**	3 (0.7%)
	Normal *n* (%)	122 (26.3%)
	Overweight *n* (%)	249 (53.7%)
	Obese *n* (%)	90 (19.4%)
Race	**White *n* (%)**	456 (98.7%)
	Black *n* (%)	6 (1.3%)
	Other *n* (%)	2 (0.4%)
Smoking Status	**Never *n* (%)**	138 (29.9%)
	Regular Smoker *n* (%)	19 (4.1%)
	Former Smoker *n* (%)	307 (66.5%)
Asthma	**Yes *n* (%)**	15 (3.2%)
	Previous *n* (%)	16 (3.5%)
	No *n* (%)	433 (93.7%)
Forced Expiratory Volume in One Second (FEV_1, L)	**mean [range]**	3.39 [1.42, 5.85]
Forced Vital Capacity (FVC, L)	**mean [range]**	2.51 [0.80, 4.32]
FEV_1/FVC ratio	**mean [range]**	73.7% [36.1%, 92.5%]
Fresh Blood Lead (µg/mL)	**mean [range]**	3.46 [0.00, 29.00]
Second Blood Sample Available	**Yes n (%)**	169 (36.4%)
Third Blood Sample Available	**Yes n (%)**	28 (6.0%)

These 661 samples were shipped to Metabolon Inc. for untargeted ultrahigh performance liquid chromatography-tandem mass spectroscopy (UPLC-MS/MS) using four platforms covering a broad range of the metabolome. Metabolites were analyzed as measured LC-MS peak areas. We then applied our in-house data processing and QC platform; metabolites with a signal-to-noise ratio <10 were considered unquantifiable and excluded, as were metabolites with undetectable/missing levels for >10% of the samples. All remaining missing values were imputed with the half the minimum peak intensity for that metabolite across the whole population, then data were *pareto* scaled to account for the differences in the scales of measurements across the metabolome, and log-transformed to create approximately Gaussian distributions and to stabilize variance. Metabolites were identified by their mass-to-charge ratio

(m/z), retention time (rt), and through a comparison to library entries of purified known standards.

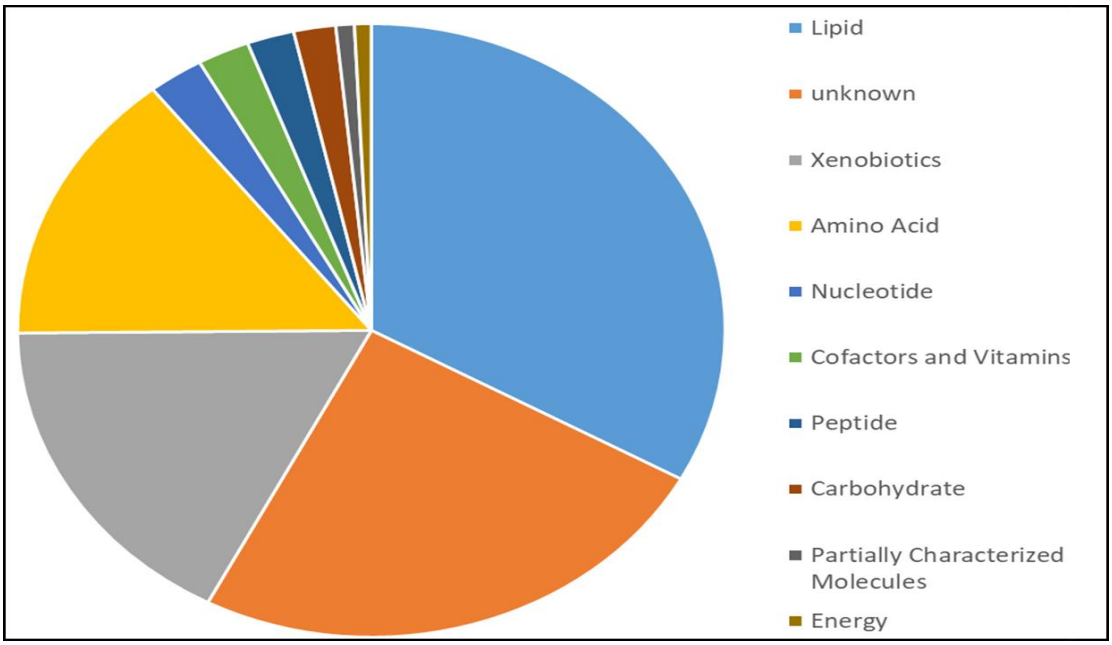
In total, relative abundance of 1301 metabolites could be quantified in the 661 plasma samples (**Table 3**). The metabolic super pathways, as defined by metabolon, covered by these profiles are described in **Figure 1**.

Table 3: Four Profiling Platforms Employed by Metabolon to Characterize 1301 Metabolites

Profiling Platform	Metabolites	n
LC/MS Negative	Metabolites that Ionize in the Negative Mode	688
LC/MS Polar	Polar Metabolites	70
LC/MS Positive Early	Metabolites that Ionize in the Positive Mode, elute early	286
LC/MS Positive Late	Metabolites that Ionize in the Positive Mode, elute late	257

LC/MS – Liquid Chromatography Mass Spectrometry

Figure 1: Metabolic Pathways Encompassed by the Metabolomic Profiles



AIM ONE: Major Task 2 Identification of the metabolome of Pb exposure

Data analyses are ongoing; in particular we are still in the process of obtaining information on longitudinal bone and toenail Pb measures, which we will use to create a composite measure of life-time Pb exposure (subtask one). All presented analyses in this report are based on Pb levels measured in blood only. We had multiple measures of blood Pb measured concurrently to the spirometry measures described in *Major task 3*.

In order to account for multiple measures per subject for Pb levels, we used a linear mixed model included race, smoking status, and BMI as fixed effects (subtask two). We considered 1301 models for each metabolite fit separately. A total of 242 metabolites (18.6%) were significantly associated with blood Pb levels. Of these 35 (2.7% were robust to False Discovery Rate (FDR) Correction according to the Benjamini and Hochberg Procedure (**Table 4**). It should be noted that there are currently no consensus standards for multiple testing correction in metabolomics; methods applied to other 'omic' datatypes such as the Bonferroni correction, and even more liberal corrections, are considered too stringent for metabolomics data due to the high correlation of metabolites that are closely linked together through biological pathways. Therefore, for the purposes of these exploratory analyses we are applying a nominal ($p<0.05$) threshold and a more conservative FDR throughout.

All analyses were conducted in R version 3.5.0 and all statistical tests were two-sided.

Table 4: Metabolites significantly associated with measures of lung function and of concurrently measured blood Pb levels

	FEV1 (L)		FVC (L)		FEV1/FVC (%)		Fresh Blood Lead (µg/mL)	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**
p<0.05	368	28.3%	189	14.5%	425	32.7%	242	18.6%
FDR corrected p<0.05	182	14.0%	4	0.3%	208	16.0%	35	2.7%

The top hits were sphingomyelin (d18:2/16:0, d18:1/16:1) ($p=2.2x10^{-6}$); Tartarate ($p=3.1x10^{-6}$); and N-acetylglycine ($p=5.1x10^{-6}$). Pathway analysis was performed with MetaboAnalyst v.4.0 (www.metaboanalyst.ca) to identify the KEGG defined (www.genome.jp/kegg/pathway.html) metabolomic pathways the 242 metabolites were enriched for. Pathway analysis extends and enhances the concept of metabolite set enrichment analysis by incorporating topology analysis. This evaluates the importance of a given metabolite based on its position within a pathway using graph theory, and therefore provides a more meaningful interpretation of list of differential metabolites. In this analysis, the hypergeometric test was specified for the over-representation analysis and relative-betweenness centrality was specified for the pathway topology analysis; all 1303 metabolites were input as the reference metabolome.

The significant metabolites were enriched for 15 metabolic pathways (**Table 5**). Among these 15 pathways, were a number that we hypothesized would be dysregulated with Pb exposure in our original grant application, including pyrimidine metabolism and alanine, aspartate and glutamate metabolism. In addition, sphingolipid metabolism, which plays a crucial role in multiple cellular functions as well as in lung health, was also identified.

Table 5: Metabolomic Pathway analysis of metabolites identified as being significantly dysregulated by degree of lung function and by blood Pb levels measured concurrently
A red square indicates that a pathway was significantly enriched among the metabolites identified as being significantly associated with the phenotype. It can therefore by hypothesized that this pathway is dysregulated with changes in the relevant phenotype

Metabolic Pathway	FEV1 (L)	FVC (L)	FEV1/FVC (%)	Fresh Blood Lead (µg/mL)
Alanine, aspartate and glutamate metabolism				■
Aminoacyl-tRNA biosynthesis	■	■	■	■
Arginine and proline metabolism	■	■		■
beta-Alanine metabolism				■
Caffeine metabolism			■	■
Cysteine and methionine metabolism		■		■
Galactose Metabolism	■			
Glycerophospholipid metabolism	■	■	■	■
Glycine, serine and threonine metabolism	■	■	■	■
Histidine metabolism	■	■		■
Lysine Degradation		■		
Nitrogen metabolism		■		■
Pantothenate and CoA biosynthesis				■
Purine Metabolism		■		
Pyrimidine metabolism		■		■
Sphingolipid metabolism		■	■	■
Taurine and hypotaurine metabolism		■	■	■
Valine, leucine and isoleucine biosynthesis	■	■	■	■

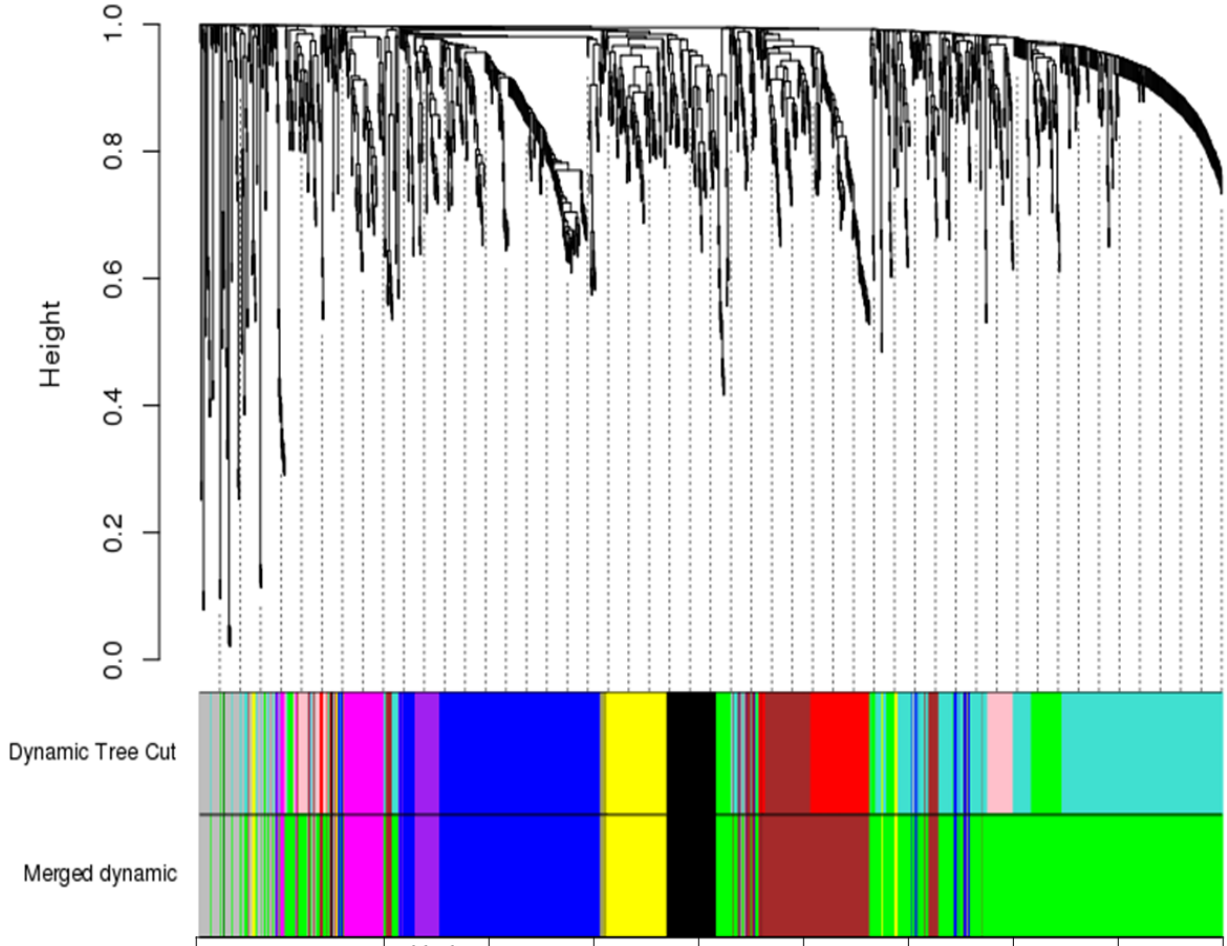
FEV1 – measure of how much air can be exhaled in one second following a deep inhalation
FVC - measurement of lung size (in liters) that represents the volume of air in the lungs that can be exhaled following a deep inhalation
FEV1/FVC – represents the percent of the lung size (FVC) that can be exhaled in one second

We then employed a network approach to identify metabolic networks, rather than single metabolites associated with Pb exposure (subtask 3). Network approaches move away from reductionist methodologies to combine systems biology and network science, providing a holistic methodology to better understand biology through the identification and investigation of non-linear relationships and networks of interacting components. Weighted Gene Correlation Network Analysis (WGCNA, horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/index) is a network method for identifying clusters or 'modules' or highly correlated variables (such as metabolites) that are likely to be co-regulated, or working together in biologically coherent fashion. A module can then be summarized as a single unit, which can be correlated with phenotypes of interest. WGCNA was used to identify metabolomic network modules within the baseline samples based on correlation patterns. The correlation matrix quantifies interconnectedness between metabolites

and assigns them to co-expression modules. Features that did not show high enough co-expression metrics with any module were excluded from further analysis (they are assigned to a redundant grey module). Highly correlated modules were then merged using a cut height (i.e. the *Euclidean* distance between clusters) of 0.6; chosen using an iterative process to identify an optimal number of adequately sized modules for analysis. Modules were summarized by an eigenvector (based on the first principal component of each module) for each participant, then associations between the modules and Pb exposures were explored.

The module assignments for the 1301 metabolites as measured in 464 samples are shown in **Figure 2.** After merging the highly correlated clusters there were eight modules, which are assigned colors by the package by default.

Figure 2: Cluster Dendrogram showing assignment of 1301 metabolites to eight merged dynamic modules



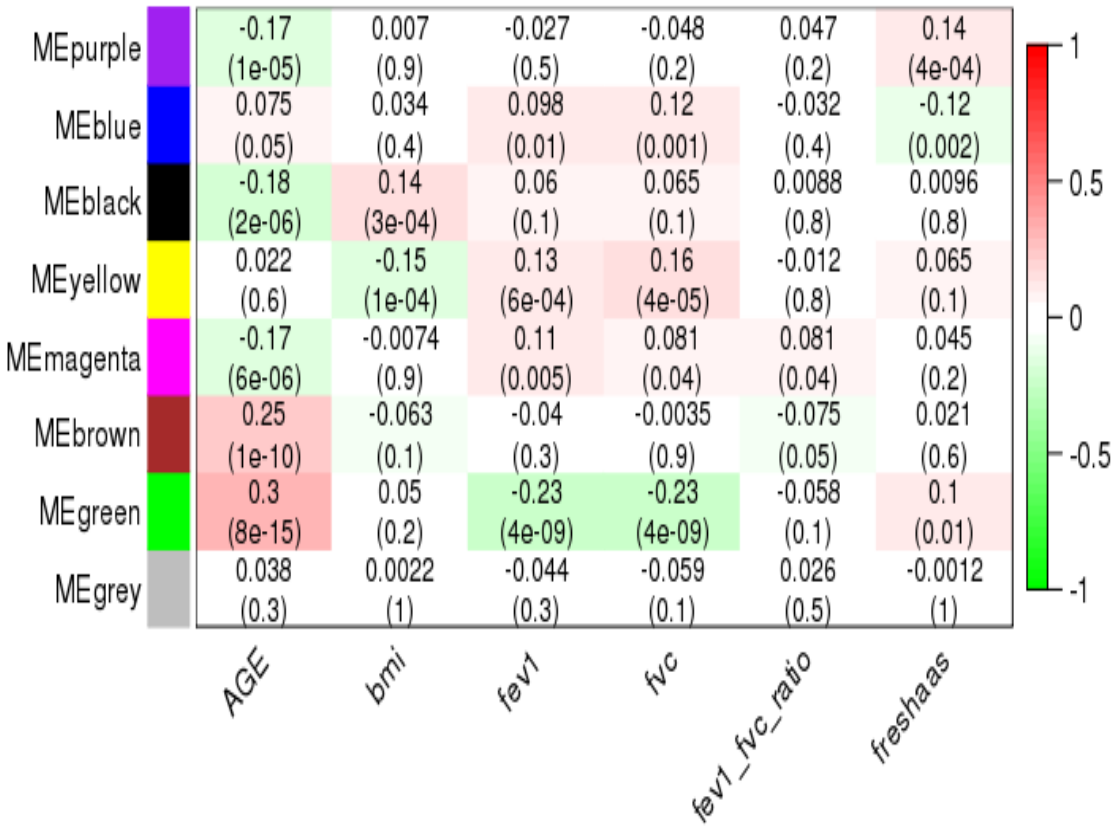
Module	black	blue	brown	green	*grey*	magenta	purple	yellow
n metabolites	69	250	193	521	*82*	60	33	93

NB. The grey module includes all those metabolites that did not show high enough correlation with any other metabolite to be included in a module. This grey module is redundant and is excluded from the analysis

Three of these modules; the purple, the blue, and the green, were significantly correlated with blood Pb (defined as freshaas in **Figure 3**). For the modules significantly ($p \leq 0.05$) correlated

with Pb levels, the 'hubs' were identified. Hubs are the features that are most highly connected within a module, and therefore drive module formation. WGCNA computes a module-membership value and associated p-value for each feature within a module, which is a measure of how connected or co-expressed that feature is with others within the same module. Features with a module-membership p-value that retained significance after Bonferroni-correction were considered to be hubs. For the purple, blue and green modules there were 31, 218 and 426 distinct hub metabolites, respectively (*see AIM TWO: Major Task one, and AIM THREE: Major Task one for further network results*).

Figure 3: Heatmap describing the relationships between the module eigengenes and the traits of interest



AIM TWO: Major Task 1: To identify a metabolomic profile of pulmonary function within a population of Veterans

Lung function was assessed according to spirometry which is performed by deeply inhaling and forcefully exhaling into a spirometer. We used the three most commonly employed measures of spirometry; (i) forced vital capacity (**FVC**);a measurement of lung size (in liters) that represents the volume of air in the lungs that can be exhaled following a deep inhalation; (ii) forced expiratory volume-one second (**FEV_1**)**;** a measure of how much air can be exhaled in one second

following a deep inhalation; and (iii) **FEV_1/ FVC** ratio which represents the percent of the lung size (FVC) that can be exhaled in one second. As we also had multiple longitudinal spirometric measures from the same men, we again employed a mixed model including race, smoking status, and BMI as fixed effects (subtask 1). The results are shown in **Table 4**. There were 368 (28.3%) metabolites associated with FEV_1, 189 (14.5%) associated with FVC and 425 (32.7%) associated with FEV1/FVC ratio. For both FEV_1 and FEV_1/FVC ratio, a large proportion of these metabolites were robust to FDR correction. There was substantial consistency in the metabolites identified in the significant end-points as would be expected given the close relationships between these outcomes (*more information on the top hits of most interest for these outcomes is provided in the AIM THREE section*). The significantly enriched pathways, which can be hypothesized to be dysregulated with changes in FEV_1, FVC and FEV_1/FVC ratio are shown in **Table 5**. Again, similar pathways were identified for the three outcomes; Aminoacyl-tRNA biosynthesis; glycerophospholipid metabolism; Glycine, Serine and Threonine metabolism; Histidine metabolism; and Valine, leucine and isoleucine biosynthesis were all significant across the outcomes, suggesting dysregulation in the pathways may be a cause or an effect of decreased lung function.

We utilized the same metabolites modules generated using an unsupervised approach in *AIM ONE: Major Task 2; subtask 3,* (**Figure 2**) and determined the relationship between these metabolite modules and the three measures of lung function (**Figure 3**). In terms of networks of interacting metabolites, there was greater concordance between FEV_1 and FVC; both were significantly correlated with the same four modules; blue, yellow, green and magenta. Interestingly, FEV_1/FVC ratio was only correlated with magenta, and additionally with the brown module. Interrogation of the 159 hub metabolites driving the formation of this brown module, determined it was characterized by Linoleic acid metabolism and fatty acid biosynthesis. These pathways have been consistently linked with lung function in the literature due to their crucial role in the mediation of the pro-inflammatory versus pro-resolving response to lung injury. Consequently, further work is required to understand why these pathways appeared to be specifically associated with FEV_1/FVC ratio, but not with FEV_1 or FVC alone. (*Further details on the network results are presented in the AIM three section*).

We next wanted to determine whether the metabolites identified in subtask 1, could be utilized to generate a metabolomics score that can be used to discriminate men by their degree of lung function (subtask 2), with the eventual aim of supporting downstream metabolite biomarker generation. We used FEV_1/FVC ratio, as a proof of principal for initial analyses to assess the optimal way to generate and test a metabolite score, and we will then utilize the optimal method to generate metabolite scores for other measures of lung function and of lead exposure. In a clinical setting FVC, FEV_1 and FEV_1/FVC ratio are compared to reference values based on healthy individuals with normal lung function, to determine the degree of lung function of the patient. The normal value for the FEV_1/FVC ratio is 70% or above, with a lower measured value corresponding to a more severe lung abnormality. In this population, 93 (20%) men had a ratio $\leq$70%, while 371 (80%) had a ratio >70%.

We compared two different models for the discrimination of an FEV_1/FVC ratio above and below 70% using receiver operator characteristic (ROC) cures and the corresponding area under the curves (AUC); Model 1: a summary score, generated by taking the first principal component

of the 189 plasma metabolites that were significantly associated with FEV_1/FVC ratio (**Table 4**); and Model 2: Levels of the four metabolites that were robust to FDR correction. The sensitivity and specificity were computed based on the optimal cut-off to maximize sensitivity and specificity weighting both equally, as determined using the 'ROCR' package in R

The summary score Model 1 had moderate discriminatory ability in a Receiver Operator Characteristic (ROC) curve analysis (AUC: 0.602 (95% CI 0.552, 0.652) while the Model 2 based on four metabolites demonstrated a marginal, and non-significant, improvement in the AUC (0.631 (95% CI 0.581, 0.680). Similarly the sensitivity was identical for the two models (55.3%) but the specificity was slightly higher for model 2 (**Table 6**). Although encouraging, these findings demonstrate the need to develop alternative methods of generating metabolite biomarkers of lung function. The resources of this project are ideally placed to do so.

Figure 4: Receiver Operator Characteristic (ROC) Curves comparing two models for the prediction of an FEV_1/FVC ratio below 70%

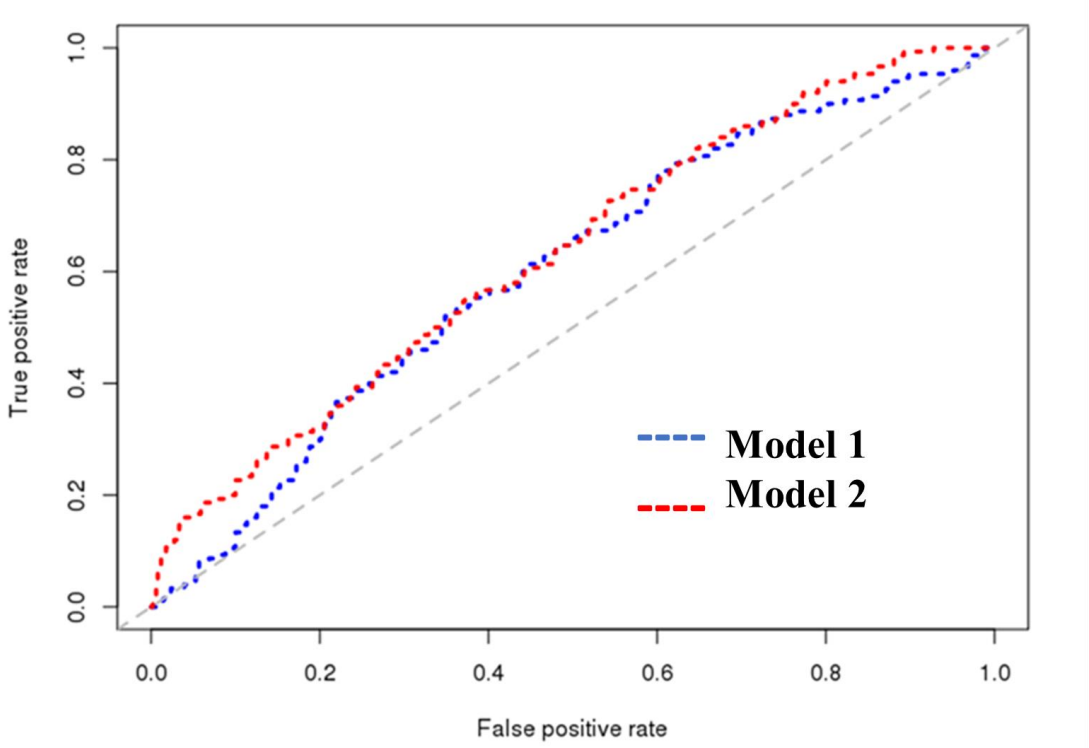


Table 6: AUCs, Sensitivity and Specificity for two models for the prediction of an FEV_1/FVC ratio below 70%

Classifier	AUC (95% CI)	Performance compared to Model 1	Sensitivity	Specificity
Model 1: Metabolite Summary Score	0.602 (0.552, 0.652)		55.3%	62.0%
Model 2: Metabolite levels	0.631 (0.581, 0.680)	*p=0.312*	55.3%	62.8%

AIM THREE: Major Task 1: Identify metabolites, metabolite networks and metabolomics pathways along the causal pathway from Pb exposure to reduced pulmonary function

To date, only subtask 1 has been started. Using the findings from Aims 1 and 2, we compared the individual metabolites and metabolite profiles (based on the WGCNA generated modules) that associated with both Pb exposure and lung function, and which may therefore lie along the causal pathway. **Figure 5**, shows the crossover between the metabolites identified as significant for FEV_1, FVC and blood Pb exposure levels based on the mixed models (**Table 4**). In total, 120 metabolites were significantly associated with at least one measure of lung function and with Pb; 72 metabolites were associated with all three indices. Some of the most biologically interesting metabolites for FEV_1 among these 72 are shown in **Table 7**.

Figure 5: Venn diagram showing the crossover between significant metabolites for two measures of lung function and for blood measures of Pb

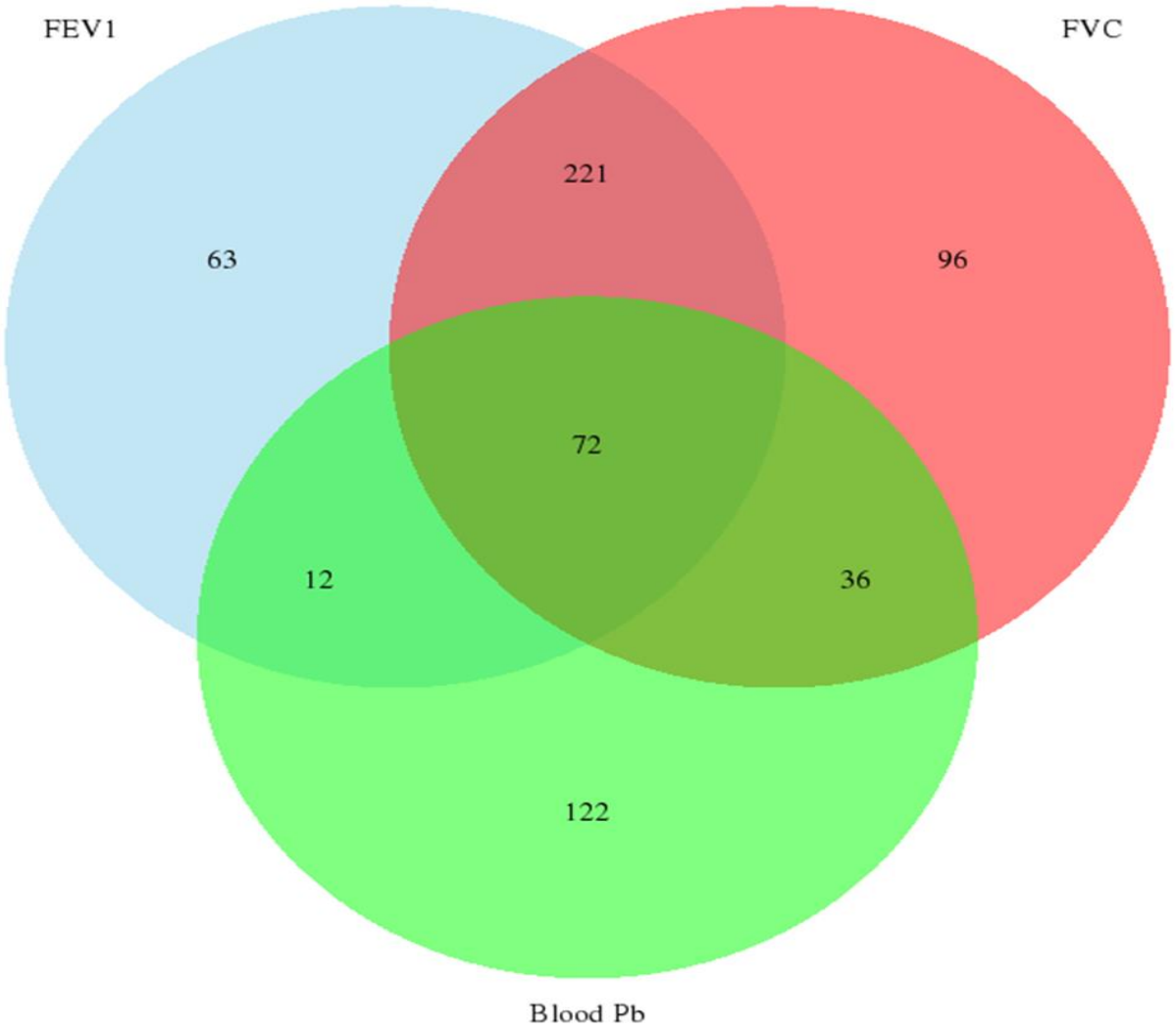


Table 6: Selection of the top metabolite hits associated with FEV1, FVC and with blood Pb

METABOLITE	SUPER PATHWAY	SUB PATHWAY	FEV1 (L)		Fresh Blood Lead (µg/mL)	
			β	p-value	β	p-value
hydroxyasparagine**	Amino Acid	Alanine and Aspartate Metabolism	-0.25	$2.5x10^{-4}$	1.58	$3.8x10^{-5}$
N-acetylalanine	Amino Acid	Alanine and Aspartate Metabolism	-0.28	0.001	1.74	$1.3x10^{-4}$
N-acetylserine	Amino Acid	Glycine, Serine and Threonine Metabolism	-0.28	$7.9x10^{-5}$	1.38	$3.4x10^{-4}$
serine	Amino Acid	Glycine, Serine and Threonine Metabolism	0.23	0.005	-1.12	0.010
N-acetylthreonine	Amino Acid	Glycine, Serine and Threonine Metabolism	-0.18	0.007	0.93	0.014
histidine	Amino Acid	Histidine Metabolism	0.23	0.014	-1.63	0.001
1-methylhistidine	Amino Acid	Histidine Metabolism	-0.12	0.031	0.62	0.034
leucine	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.31	0.002	-1.61	0.001
alpha-hydroxyisocaproate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.17	$8.5x10^{-5}$	-0.61	0.008
choline phosphate	Lipid	Phospholipid Metabolism	0.11	0.019	-0.93	$1.3x10^{-4}$
phosphoethanolamine	Lipid	Phospholipid Metabolism	0.11	0.007	-0.76	$2.5x10^{-4}$
sphingadienine	Lipid	Sphingolipid Synthesis	0.08	0.005	-0.48	0.002
sphinganine	Lipid	Sphingolipid Synthesis	0.09	0.039	-0.66	0.002
sphinganine-1-phosphate	Lipid	Sphingolipid Synthesis	0.14	0.005	-0.66	0.008
sphingosine	Lipid	Sphingosines	0.09	0.020	-0.61	0.004
sphingosine 1-phosphate	Lipid	Sphingosines	0.16	0.003	-0.66	0.020

The identified metabolites were primarily amino acids and lipids, and in keeping with the results of the pathway analysis show in **Table 5**, they were specifically involved in the pathways of Alanine and Aspartate Metabolism; Glycine, Serine and Threonine Metabolism; Histidine Metabolism; Leucine, Isoleucine and Valine Metabolism; Phospholipid Metabolism and Sphingolipid synthesis. Interestingly, it was uniformly observed that metabolites that were inversely associated with lung function – i.e. those which were increased with decreased lung function, were positively associated with blood Pb. In others words, these results supported our hypothesis that increased blood levels lead to decreased lung function, and that metabolites may mediate this association and help us to underlying that biological mechanisms underlying it.

The findings from the mixed models were also supported by those from the network analysis. Two modules were associated with both lung function and with Pb; the green and the blue (**Figure 3**). We performed pathway analysis of the hub metabolites for these two modules; and showed that the blue module was enriched for pathways including Glycine, Serine and Threonine Metabolism; Alanine and Aspartate Metabolism and Sphingolipid metabolism; while the green module was enriched for histidine metabolism. This suggest network methods may be able to capture groups of interacting metabolites along that putative causal pathway, proving greater biological insights into the results. These are novel findings in a human population, which are supported by experimental evidence in the literature.

However, to robustly test such a hypothesis requires more formal mediation and structural equation analyses, as outline in the SOW, *AIM THREE; Major task two*. In keeping with the proposed timeline, *AIM THREE; Major task two* as well as *Major Task three* are in development.

The literature searches as outlined in the SOW have also been completed and will inform the manuscripts that are in development.

The initial findings have been accepted for publication at the American Thoracic Society Annual Meeting, one of the premier conferences for respiratory research, which will be held in Dallas in May 2019. They have also been submitted to Northeastern University's 2019 Research Innovation and Scholarship Expo to be held in April in Boston We anticipate the findings will also be submitted to the International Metabolomics Society Annual meeting in 2019, when abstract submission opens.

3.c. Opportunities for training and professional Development Provided by the Project

Throughout this reporting period, Dr. Lasky-Su has been closely mentoring Dr. Kelly. Due in large part to the skills that she has learned and been able to put in practice including budget management, co-leading a study, working with collaborators at the normative Ageing Study, sample selection and data analysis, Dr. Kelly has now been promoted from a post-doctoral fellow to a junior faculty member in the Channing Division of Network Medicine.

3.d. Dissemination of results to communities of interest

Multiple presentations relating to this proposal, including theoretical study design and preliminary findings have been presented to colleagues in the fields of metabolomics and respiratory health within the Channing Division of Network Medicine. The findings will be presented to the wider thoracic community at the American Thoracic Society Annual conference in May 2019, and Northeastern University's 2019 Research Innovation and Scholarship Expo. Dependent on findings we also anticipate submitting updated and novel findings to the American Thoracic Society 2020 Annual conference.

3.e. Plans for the next reporting period to accomplish these goals

During the next reporting period, we plan to present our initial findings as an oral presentation at the American Thoracic Society Annual conference (Dallas, USA) and the International Metabolomics Society International conference (The Hague, Netherlands). We then aim to publish these findings in a high-ranking peer reviewed journal, as delineated in our **SOW** (**Table1a&b**) we plan to publish a minimum of two initial manuscripts; the metabolome of Pb exposure and the metabolome of respiratory health within a veteran's population; both of which represent novel additions to the literature.

Concurrently, we aim to complete the remaining goals and their subtasks, culminating in further publications (**Table1c**) and the acquisition of further funding in order to extend this project towards the forward translation into a clinical setting, as well as reverse translation into the development of therapeutic targets (*Specific Aim 3: Major task 2)* As part of this effort we will continue to search for collaborators with suitable replication populations. Furthermore, throughout this award period we have been furthering our involvement in the International Metabolomics community; in particular Dr. Jessica Lasky-Su has been voted in as chair of the Consortium of Metabolomics Studies (COMETS) committee and to the board of the Metabolomics Society. Dr. Kelly is a steering committee member. We will continue to work with the metabolomics community to develop novel methodologies for the analysis of metabolomics data, to ensure can successfully complete our aims and drive forward the field.

4. Impact

4.a. Impact on the development of the principal discipline of the project

The principal discipline of this project is the field of metabolomics. The overarching goal is to construct a casual pathway between Pb exposure and poor respiratory health in a cohort of veterans, with the hypothesis that this pathway is mediated through measurable metabolomic pathways. To do so requires the development of novel analytical and statistical techniques with a focus on network methodology. We have, and will continue to work closely with our bioinformatics and network scientist colleagues to develop the appropriate methodologies. These techniques will then be applicable to a wide range of projects in the field of metabolomics, where the construction of causal pathways mediated by metabolites is a key, but as yet unachieved, goal for many.

4.b. Impact on other disciplines

In *Specific aim 1* (**Table 1a**), we propose to identify the metabolome of past Pb exposure that takes both duration and intensity into account. Our initial findings have been very promising in this area and, as such, support the utility of metabolomics in the disciplines of exposure science and exposure biomarker development. Exploring respiratory disease through metabolomics is more established, however these results also add to the discipline of respiratory health, though (i) the addition of novel literature in an underrepresented population, and (ii) supportive mechanistic evidence for a link between Pb exposure and poor respiratory health.

4.c. Impact on technology transfer

Nothing to report

4.d. Impact on society beyond science and technology

To date, we have demonstrated evidence to support our hypothesis of a link between Pb exposure and respiratory health, mediated through dysregulated metabolism, as outlined in our proposal, the results of this study will be of great importance to the growing population of who were exposed to Pb during active service in the Gulf, Iraq and Afghanistan, as well as to the wider US population in whom Pb exposure still widespread. The findings will support **(i) Prevention:** by confirming that Pb exposure causes reduced lung function, thereby supporting improved regulations and safeguards pertaining to exposure in the military **(ii) Early intervention** through the development of markers in the blood that can be used to identify the Pb exposure at the critical level that can damage respiratory health **(iii) Treatment of Pb induced pulmonary dysfunction** through the identification of molecules that are affected by Pb exposure and influence respiratory health, which can then be targeted by novel drugs and therapies.

5. Changes/Problems

5.a Changes in approach and reasons for change

There was a significant delay in the receipt of HRPO approval. It was initially determined that a new IRB approval was required. However, after a few months of discussion it was subsequently concluded that "the activities conducted by the Partners investigators do not constitute human subjects research". As such no new IRB approval was required, and the project could proceed. This resulted in a delay of approximately 6 months. However, as the issue was resolved, this did not ultimately results in any changes to our approach or analytical plans.

5.b. Changes that had a significant impact on expenditures

Due to price negotiations with Metabolon Inc., who performed the metabolomic profiling, we were able to profile 661 samples, rather than the anticipated 374, this increased the power and impact of our proposal. *Further details in Accomplishments. Section 3.a*

5.c. Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Nothing to report

6. Products, Inventions, Patent Applications, and/or Licenses

6.a Publications, conference papers, and presentations

Two papers currently under preparation; (i) The metabolome of Pb exposure; (ii) the metabolome of respiratory health in a cohort of veterans.

One conference paper accepted to the American Thoracic Society Annual Conference, Dallas, USA 2019.

6.b. Websites or Other Internet Sites

Nothing to report

6.c. Technologies or techniques

Novel statistical techniques under development (see *Impact section 4.a*)

6.d. Inventions, patent applications, and/or licenses

Nothing to report

6.e Other Products

Comprehensive Metabolomic Dataset. The generation of a database of 661 longitudinal plasma samples with metabolomic profiling from 464 military veterans with measures of Pb exposure and lung function, provides an invaluable research resource beyond the scope of this current project. This database will allow us to address a large number of research questions and provide a valuable validation population for an ongoing project exploring the link between the metabolome and body mass index. Further, we have been in discussion with other collaborators regarding a project exploring the metabolome of Aging. The age range and baseline characteristics of this population make them perfectly suited for such a project. We anticipate that other collaborations based around this database will follow.

7. Participants & Other Collaborating Organizations

7.a. Individuals who have worked on this project

Name:	*Jessica Lasky-Su*
Project Role:	*PI*
Researcher Identifier (e.g. ORCID ID):	*Jessica.a.su@gmail.com*
Nearest person month worked:	*3 person months*
Contribution to Project:	*Dr. Lasky-Su is the PI of this project and has overseen all aspects and in particular the management of the budget and working with Metabolon to negotiate metabolomic profiling prices. Dr. Lasky-Su is now working directly with Dr. Kelly on the data analysis and the preparation of manuscripts*
Funding Support:	*This DOD award*

Name:	*Rachel Kelly*
Project Role:	*Co-PI*
Researcher Identifier (e.g. ORCID ID):	0000-0003-3023-1822
Nearest person month worked:	*3 person months*
Contribution to Project:	*Dr. Kelly is the co-PI of this project. She worked with collaborators at the Normative Ageing study to collect the samples, applied the QC and data processing to the data and is now leading the data analysis. Dr Kelly is working with Dr Lasky-Su to prepare the first manuscripts and had submitted an abstract based on this work as first author to the American Thoracic Society Annual conference*
Funding Support:	*This DOD award*

7.b. Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Nothing to report

7.c. What other organizations were involved as partners?

Nothing to Report