

Proteomic profiling reveals immunologic mediators of endotoxin-related lung function decline: a longitudinal prospective study in textile workers

Received: 29 November 2025

Accepted: 4 April 2026

Published online: 17 April 2026

Cite this article as: Wang H., Goodney G., Zhang F. *et al.* Proteomic profiling reveals immunologic mediators of endotoxin-related lung function decline: a longitudinal prospective study in textile workers. *Respir Res* (2026). <https://doi.org/10.1186/s12931-026-03669-4>

Hantao Wang, Gabriel Goodney, Fengying Zhang, Peggy S. Lai, Li Su, Peng Gao, Jungeun Lim, Shelton Lo, Zhonghua Liu, David C. Christiani & Jason Y. Y. Wong

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Proteomic Profiling Reveals Immunologic Mediators of Endotoxin-Related Lung Function Decline: A Longitudinal Prospective Study in Textile Workers

Hantao Wang¹, Gabriel Goodney¹, Fengying Zhang², Peggy S. Lai^{3,4}, Li Su⁵, Peng Gao⁵, Jungeun Lim¹, Shelton Lo¹, Zhonghua Liu⁶, David C. Christiani^{3,4,5}, Jason Y.Y. Wong^{1*}

¹ Epidemiology and Community Health Branch, National Heart, Lung, and Blood Institute, Division of Intramural Research, National Institutes of Health, 10 Center Drive, Bethesda, MD, USA, 20892

²Shanghai Putuo District People's Hospital, No. 1291, Jiangning Rd, Putuo District, Shanghai, China

³Harvard Medical School, 25 Shattuck St, Boston, MA, USA, 02115

⁴Massachusetts General Hospital, 55 Fruit Street, Boston, MA, USA, 02114

⁵Department of Environmental Health, Harvard T.H. Chan School of Public Health, 665 Huntington Avenue, Boston, MA, USA, 02115

⁶Department of Biostatistics, Mailman School of Public Health, Columbia University, 722 W 168th St, New York, NY, USA, 10032

*Corresponding Author:

Jason YY. Wong, Email: Jason.wong@nih.gov

Author Contribution:

DCC established the cohort and supervised the project. FZ, PSL, and LS collected data by administering the surveys and conducting spirometry examinations. HW and GG conducted statistical analysis. HW, JL, PG, ZL, DCC, and JYYW contributed to the conceptual design of this manuscript and interpretation of the results. HW prepared the manuscript. All authors reviewed and approved the final version of the manuscript.

Ethical approval was obtained from the Institutional Review Boards of the Harvard School of Public Health (protocol # 10175) and Shanghai Putuo District People's Hospital.

Abstract

Background: Endotoxin is a major pathogenic component of cotton dust in the textile industry. While airborne endotoxin is associated with lung function decline, the underlying molecular mechanisms remain unclear. Proteomic profiling could provide important insight into the pathways that drive the harmful respiratory effects of endotoxin.

Methods: We conducted serum proteomic profiling of 221 endotoxin-exposed cotton workers and 192 endotoxin-free silk workers from the longitudinal cohort of the Shanghai Textile Worker Study. Using blood samples collected in 2016, proteins were quantified by data-independent acquisition mass spectrometry and genotyping was assessed using low-pass whole genome sequencing. Forced expiratory volume in 1 second (FEV-1) was measured in 2011 and 2016. We used adjusted regression to identify differentially expressed proteins (DEPs) between the exposed and control groups. Causal mediation analyses were performed to identify protein mediators. Mendelian Randomization analyses provided causal estimates of protein effects on lung function change.

Results: Among 2,962 quantified proteins, we identified 224 proteins that were differentially expressed between cotton and silk workers (Bonferroni $p < 0.05$). Top enriched pathways associated with DEPs were the

complement and coagulation cascades (KEGG: hsa04610) and the chemokine signaling pathway (KEGG: hsa04062). Adaptive immune proteins collectively mediated endotoxin-related lung function decline (binomial $p = 9.63 \times 10^{-97}$). Two immunoglobulin domain proteins significantly (Bonferroni $p < 0.05$) mediated endotoxin effects on lung function change: Epididymis luminal protein 180 (HEL180) and IGL c3728_light_IGKV4-1_IGKJ1 (IGL c3728), with 46.4% ($p = 0.04$) and 46.9% ($p = 0.05$) proportion of the total effect that was mediated. MR estimates demonstrated that every 2-fold decrease in HEL180 and IGL c3728 expression was associated with an FEV-1 decline of 1.90 ml/year (95% CI: [0.79, 3.02]) and 2.36 ml/year (95% CI: [1.68, 3.02]), respectively.

Conclusion: Using a trans-omic approach, our findings suggest that chronic endotoxin exposure suppresses immunoglobulin domain proteins, weakens adaptive immunity, and accelerates lung function decline. These findings provide greater precision in understanding biological mechanisms underlying endotoxin-related respiratory dysfunction.

Keywords: Cotton dust, occupational health, occupational epidemiology, vegetable dust.

Background

Textile manufacturing is one of the largest traditional industries worldwide employing approximately 430 million people¹. Since the 1970s, textile manufacturing has increasingly shifted to low- and middle-income countries (LMICs) as the rise of "fast fashion", characterized by rapid production cycles and the frequent new collections, has intensified production demands². Given that cotton is the most commonly used fiber in the textile industry, identifying and mitigating occupational health and safety risks is critically important, particularly in LMICs with less stringent labor regulations. As reported by the Economic Research Service of the U.S. Department of Agriculture, about 25% of all textile products are made of cotton. Cotton textile production involves organic dust, particularly during yarn preparation and processing. Textile workers encountering organic dust with elevated levels of bacterial endotoxin are at greater risk of respiratory conditions.

The primary toxicant in cotton dust is endotoxin or lipopolysaccharide (LPS), which is often found in the outer membrane of gram-negative bacteria. Endotoxin signals through Toll-like receptor 4 (TLR4) and triggers signaling pathways that involve the adaptor protein MyD88. This leads to NF- κ B activation and the production of pro-inflammatory cytokines³⁻⁵. Our previous studies on textile workers in Shanghai, China⁶⁻¹¹ have linked endotoxin exposure to lung function decline. The association between

endotoxin and adverse health effects is consistent with findings from other occupational studies of textile workers conducted worldwide¹², especially in LMICs¹³. These studies suggest that endotoxin triggers airway inflammation and hyper-responsiveness, ultimately leading to airway obstruction.

Occupational exposure to endotoxin is not unique to textile workers.

Individuals in other workplaces with similar organic dust exposures, such as animal feeding operations¹⁴, are also highly exposed to endotoxin. Although our previous studies have examined how protein expression affects lung function¹⁵, the molecular mechanisms by which proteomic features mediate the effects of chronic endotoxin exposure on lung function decline remain unclear.

In this study, we first identified biological pathways involved in chronic endotoxin exposure by comparing protein expression between endotoxin-exposed cotton workers and endotoxin-free silk workers. Next, we investigated how specific protein features causally mediated the effect of endotoxin on changes in lung function. Finally, by integrating genomic and proteomics data, we estimated how changes in protein mediator expression levels affected lung function.

Methods

Study Design and Study Population

The Shanghai Textile Worker Study is a longitudinal prospective cohort study initiated in 1981, consisting of 447 cotton workers and 472 silk workers residing in Shanghai, China. Questionnaires on work history, respiratory symptoms, smoking history, and demographics were collected every five years along with spirometry tests. Details of the cohort are outlined in our supplementary materials and have been reported previously⁶⁻¹¹. Ethical approval was obtained from the Institutional Review Boards of the Harvard School of Public Health and Shanghai Putuo District People's Hospital.

Exposure Assessment of Endotoxin

Area full-shift samples of cotton dust were collected on-site during each survey until the mill closures in 2001. Endotoxin concentrations were analyzed using the Limulus Amebocyte Lysate assay⁶⁻¹¹. The cumulative endotoxin concentration (EU/m³) was approximated using job titles and work history.

Respiratory Outcome Assessment

Forced expiratory volume in 1 second (FEV-1) was measured according to the spirometry technical standards of the American Thoracic Society (ATS)¹⁶. The primary outcome was the annual change in FEV-1 between the 2011 and 2016 surveys (ml/year), with positive values indicating an improvement in lung function and negative values indicating a decline.

Proteomics

Serum samples were collected from 221 endotoxin-exposed cotton workers and 216 endotoxin-free silk workers in 2016. Protein quantification was performed using data-independent acquisition mass spectrometry (DIA-MS). Details of sample preparation can be found in our previous study¹⁵ and in the supplementary materials. Protein expression levels were $\log_2$ transformed and normalized. Proteins with more than 50% missing values were excluded from the analysis. Missing values were imputed using multiple imputation by chained equations (MICE). Each protein was assigned to one of five immune types (adaptive, innate, both, unspecified, or non-immune), based on whether its Gene Ontology (GO)¹⁷ biological process annotations included any child term of the adaptive immune response (GO:0002250), innate immune response (GO:0045087), or the broader immune system process (GO:0002376).

Low-pass Whole Genome Sequencing

Genotyping was performed by BGI (formally Beijing Genomics Institute) using low-pass whole genome sequencing (WGS) and imputed with Gencove's ImputeSeq¹⁸, with the same blood samples collected for proteomics. Single Nucleotide Polymorphisms (SNPs) were filtered by excluding the X chromosome, having a minor allele frequency (MAF) <

0.05, failing Hardy-Weinberg equilibrium (HWE) at $p < 1 \times 10^{-6}$, and removing duplicated Reference SNP cluster IDs (rsid).

Statistical Analysis

Differentially Expressed Protein Analysis: Four models were fit to compare protein expression levels between cotton workers and endotoxin-free silk workers. Each model assessed the association using either raw or MICE-imputed data. Endotoxin exposure was evaluated either as a binary variable (cotton vs. silk) or as a continuous variable (natural log-transformed cumulative endotoxin concentration). All models adjusted for the log₂-transformed protein expression levels for sex (male, female), age (years), smoking pack-years, body mass index (BMI; kg/m²), and the first three principal components (PCs) of all protein expressions. Differentially expressed protein (DEP) candidates were determined when proteins were significant in at least two of the four regression models. Bonferroni-corrected p-values < 0.05 were applied to control for multiple testing. To further evaluate whether immune-related proteins were collectively significant ($p < 0.05$), binomial enrichment tests were conducted separately for innate and adaptive immune proteins.

Correlation Network Analysis: A global protein-protein correlation network was first constructed using Prim's maximum spanning tree (MST)²⁰ to characterize overall protein correlation patterns among all participants.

Occupation-stratified networks of identified DEPs were then constructed to compare differences between cotton and silk workers. To further understand how immune-related proteins interacted within each occupation, immune-specific subnetworks were extracted and analyzed. All networks were evaluated both visually and quantitatively using topological metrics, including network diameter, radius, and heterogeneity. More flexible network structures were generated using Triangulated Maximally Filtered Graph (TMFG)²¹ as sensitivity analysis.

Functional Enrichment: Enrichment analysis was performed using gene annotation from STRING¹⁸. A signal score, defined as the harmonic mean of the false discovery rate (FDR) and the proportion of observed genes to background genes, was used to quantify the strength of enrichment. For each biological term or pathway, the proportion of downregulated genes was determined, with the statistical significance evaluated by the Wilcoxon signed-rank test.

Causal Mediation Analysis: High-Dimensional Mediation Testing (HDMT)²² and the modified Divide-Aggregate Composite-null Test (DACT)^{23,24} were applied to identify proteins that mediated the association between endotoxin exposure and lung function decline. Both mediation and outcome models were adjusted for sex, age, smoking pack-years, height, BMI, and the first three PCs of overall protein expression. Bonferroni-corrected p-

values < 0.05 were applied to control for multiple testing. The significance of the proportion of the association mediated by each significant protein mediator was estimated using bootstrapping ($n=1000$). Binomial enrichment tests were conducted to determine whether innate and adaptive immune proteins were significantly ($p<0.05$) mediated endotoxin-related lung function decline.

Mendelian Randomization: Significant protein mediators identified through causal mediation analysis were integrated with low-pass GWAS data to identify SNPs associated with these protein mediators using an additive model. To reduce horizontal pleiotropy, SNPs were filtered by linkage disequilibrium (LD) pruning, retaining only SNPs with $r^2 < 0.5$ within 50 base pairs. Candidate SNPs were then used as instrumental variables to assess the causal effect of protein expression on lung function decline. Inverse Variance Weighting (IVW) was employed as the primary Mendelian Randomization (MR) method with other commonly used MR methods as sensitivity analyses. SNPs with p -values $< 5 \times 10^{-6}$ for both the SNP-protein association and the SNP-lung function mediation test using DACT were selected for MR analysis. In addition, the effects of protein expression on lung function decline were estimated directly using adjusted ordinary least squares (OLS) regression models for comparison.

Results

Characteristics of Participants

The participant selection and statistical analysis steps are listed in Figure 1. The characteristics of the study population are shown in Table 1. Baseline characteristics of cotton and silk workers were similar, with no significant differences in age, sex, and smoking status distribution ($p>0.05$). Cotton workers had sharper annual declines in FEV-1 compared to silk workers between 2011 and 2016 (-25 vs -23 ml/year, $p=0.02$). Cotton workers lost more in FEV-1 over the 35-year follow-up compared to silk workers (900 ml or 30% vs 800 ml or 27.6%).

DEPs between Endotoxin Exposed Cotton and Endotoxin Free Silk Workers

A total of 2,962 proteins were quantified. Out of the 1,233 protein features with less than 50% missing data, 592 proteins were annotated as immune-related proteins (95 innate and 144 adaptive immune proteins). Differential expression analysis revealed that 224 proteins were differentially expressed between endotoxin-exposed cotton and endotoxin-free silk workers. Of these 224 Bonferroni-significant DEPs, 96 proteins were immune-related proteins (20 innate and 40 adaptive immune proteins). Both adaptive (binomial $p=9.63 \times 10^{-97}$) and innate (binomial $p=6.59 \times 10^{-33}$) immune protein expressions were collectively disrupted by endotoxin, as indicated by highly enriched proportions, suggesting broad immune downregulation with chronic endotoxin exposure. Adaptive and innate proteins were significant

(binomial $p < 0.05$) in sub-cohort analyses of female, smoker, non-smoker and cotton worker only. (Supplement Table S1)

Correlation Networks Analysis

In the MST constructed for all workers (Supplement Figure S1), Profilin-1 (UniProt ID: P07737; betweenness centrality = 0.69) was identified as the central node. Elongation factor G, mitochondrial (UniProt ID: Q96RP9) exhibited the highest degree within the MST, particularly connecting with a large number of proteins with limited annotation.

When comparing the stratified networks by occupation, the network of cotton workers (Figure 2B) displayed a more complex and heterogeneous structure than that of the silk workers (Figure 2A). Cotton workers compared to silk workers showed higher network diameter (30 vs 25), network radius (15 vs 13), and network heterogeneity (0.90 vs 0.82). The more heterogeneous network structure in cotton workers was also observed in immune subnetworks and sensitivity graphs (Supplemental Table S2 and Supplemental Figure S2). These topological differences suggest that endotoxin exposure may disrupt protein expression and alter the protein-protein correlation structure, particularly in immune-related proteins by increasing variability and forming more extreme network hubs.

Function Enrichment of DEPs

For cellular components, we observed significant differences in protein expression levels, particularly among vesicle- and extracellular region-related components. Cotton workers exhibited significantly altered protein expression in vesicles (GO:0031982, $n = 94$, $FDR = 1.49 \times 10^{-32}$, signal = 1.48) and extracellular spaces (GO:0005615, $n = 99$, $FDR = 2.37 \times 10^{-44}$, signal = 1.92) compared to silk workers, especially within blood microparticles (GO:0072562, $n = 18$, $FDR = 3.54 \times 10^{-16}$, signal = 3.18). Among vesicle-related components, both intracellular vesicles (GO:0097708, $n = 49$, $FDR = 2.10 \times 10^{-10}$, signal = 0.96) and extracellular exosomes (GO:0070062, $n = 80$, $FDR = 2.04 \times 10^{-39}$, signal = 2.26) were significantly enriched. The most enriched component within intracellular vesicles was the secretory granule lumen (GO:0034774, $n = 27$, $FDR = 1.23 \times 10^{-18}$, signal = 2.82).

At the level of molecular function, proteins associated with hydrolase activity (GO:0016787) were significantly downregulated in cotton workers. This included serine-type endopeptidase inhibitor activity (GO:0004867, $n = 15$, $FDR = 1.95 \times 10^{-8}$, signal = 1.2) and GTPase activity (GO:0003924, $n = 11$, $FDR = 0.0055$, signal = 0.51). Additionally, proteins involved in binding functions such as cadherin binding (GO:0045296, $n = 13$, $FDR = 0.00003$, signal = 0.75) and GTP binding (GO:0005525, $n = 12$, $FDR = 0.0055$, signal = 0.51) were significantly less expressed in the cotton group.

In terms of biological processes, 135 distinct processes showed differential enrichment between cotton and silk workers. The processes with the highest signal values included regulation of endopeptidase activity (GO:0052548, $n = 20$, $FDR = 1.20 \times 10^{-7}$, $signal = 1.28$), inflammatory response (GO:0006954, $n = 20$, $FDR = 2.88 \times 10^{-5}$, $signal = 1.01$), and negative regulation of hydrolase activity (GO:0051346, $n = 16$, $FDR = 8.48 \times 10^{-6}$, $signal = 1.01$). Regulation of the immune system process was also significantly enriched (GO:0002682, $n = 29$, $FDR = 8.95 \times 10^{-5}$, $signal = 0.65$).

KEGG pathway analysis revealed significant enrichment in both the complement and coagulation cascades (KEGG: hsa04610, $n = 9$, $FDR = 3.94 \times 10^{-6}$, $signal = 1.30$) and the chemokine signaling pathway (KEGG: hsa04062, $n = 11$, $FDR = 8.25 \times 10^{-5}$, $signal = 1.04$). The top three enriched Reactome pathways were platelet activation, signaling and aggregation (HSA-76002, $n = 31$, $FDR = 5.41 \times 10^{-25}$, $signal = 3.90$), platelet degranulation (HSA-114608, $n = 21$, $FDR = 3.29 \times 10^{-19}$, $signal = 3.71$), and hemostasis (HSA-109582, $n = 36$, $FDR = 4.09 \times 10^{-20}$, $signal = 2.46$).

On average, protein expression levels were significantly (Wilcoxon signed-rank test $p < 0.05$) lower in cotton workers compared to silk workers across

all biological terms and pathways, with a minimum proportion of downregulated genes of 78.57%. Detailed enrichment results are shown in Supplemental Table S3.

Protein Mediators of Endotoxin-Related Lung Function Change

Causal mediation analysis was conducted to further identify proteins that significantly mediate the effect of endotoxin exposure on lung function decline among the noteworthy DEPs. Among the 224 endotoxin-associated DEPs, 18 immune proteins showed significant (HDMT $p < 0.05$) mediation of the endotoxin-lung function decline relationship (binomial $p = 3.65 \times 10^{-7}$). Enrichment was driven by adaptive immune proteins ($n=12$, binomial $p = 1.21 \times 10^{-6}$), whereas innate immune proteins showed no enrichment ($n=1$, binomial $p = 0.64$). These results indicate that adaptive immune pathways are the predominant mediators linking endotoxin exposure to lung function decline, while innate components play a lesser role.

In addition, two immunoglobulin (Ig) domain proteins, Epididymis luminal protein 180 (Uniprot ID: B6EDE2, HEL180) and IGL c3728_light_IGKV4-1_IGKJ1 (Uniprot ID: A0A5C2G374, IGL c3728), were identified as significant mediators (Bonferroni $p < 0.05$) in models using both raw and imputed protein expression and using both continuous and binary exposure terms. Results were significant in both HDMT and modified-DACT testing (Table 2). After 1,000 bootstrapping iterations, the proportion mediated was

estimated at 46.4% ($p = 0.04$) for HEL180 and 46.9% ($p = 0.05$) for IGL c3728. Mediation significance of both IGL c3728 and HEL180 were significant (HDMT $p < 0.05$) in sex-stratified analysis.

Effect of Protein Expression on Lung Function Change

We analyzed a total of 6,323,974 SNPs in the mediation analysis to evaluate how significant protein mediators affect lung function change, adjusted for endotoxin exposure. Fourteen SNPs were associated with HEL180 expression ($p < 5 \times 10^{-6}$), and another fourteen SNPs were associated with IGL c3728 expression. The SNP lists were identical between models using binary and continuous endotoxin exposure terms. Details of the associated SNPs are provided in Supplemental Table S4.

A forest plot illustrating the estimated effects of HEL180 and IGL c3728 on lung function change is shown in Figure 3. In general, protein expressions of HEL180 and IGL c3728 demonstrated a positive association with lung function change. Using IVW, a two-fold decrease in HEL180 expression was associated with a 1.92 ml/year (95% CI: [0.79, 3.02]) decline in FEV-1. For IGL c3728, the decline was 2.35 ml/year (95% CI: 1.68, 3.02)). Finally, in OLS models, every two-fold decrease in HEL180 expression was associated with a 1.79 ml/year (95% CI: [0.86, 2.72]) decline in FEV-1. A similar decline in FEV-1 was observed for every two-fold decrease in IGL c3728

expression (1.64 ml/year, 95% CI: [0.71, 2.57]). Estimates from OLS models were directionally consistent with MR analyses.

Discussion

We leveraged advanced proteomic profiling, genomics, and causal methods to understand the molecular changes and biological pathways underlying endotoxin-related lung function decline. We identified 224 differentially expressed protein features between endotoxin-exposed cotton workers and endotoxin-free silk workers. Most of these proteomic features were related to inflammatory responses and immune pathways. Among these noteworthy signals, two significant Ig protein features were found to significantly mediate endotoxin-related lung function change: HEL180 and IGL c3728, suggesting endotoxin potentially affects lung function decline through the adaptive immune response.

Our correlation network and enrichment analyses revealed that chronic endotoxin exposure may disrupt both innate and adaptive immunity, leading to impairment of immune and inflammatory regulation. Most of the 224 DEPs between endotoxin-exposed and control textile workers were associated with vesicle-related cellular components. Vesicles, including extracellular vesicles (EVs), are essential for transferring molecules between cells, including immune cells. We also observed downregulation of hydrolase activity in response to endotoxin exposure. Hydrolases are often

packaged within vesicles such as lysosomes or secretory granules. Downregulation of hydrolase activity may lead to tissue damage and persistent inflammation²⁵. In addition, 78–100% of proteins were downregulated in key inflammatory and immune pathways such as the complement cascade, chemokine signaling, and platelet activation, suggesting broad suppression of immune responses. Collectively, our findings provide evidence supporting endotoxin tolerance in chronically exposed workers, where innate and adaptive immune responses are possibly reprogrammed following repeated exposure to LPS. Such alterations ultimately impair host defense and hemostatic functions that are critical for maintaining lung health²⁶.

Through mediation analysis, we identified HEL180 and IGL c3728 as significant mediators of the effect of endotoxin exposure on lung function. The protein expression of these two mediators was moderately correlated ($r = 0.40$). Both proteins belong to the Ig domain family. Ig domain proteins play key roles in adaptive immune responses, including antigen recognition and immune signaling²⁷. In the respiratory system, local antibody production is critical for the neutralization of inhaled pathogens and mucociliary clearance. While HEL180 was primarily characterized as an epithelial secretory protein in reproductive tissue, the identification of HEL180 as mediator protein may reflect a research focus on tissue-specific rather than true sex-restricted expression. Many epithelial secretory

proteins initially identified in the epididymis have since been found to be expressed in diverse mucosal tissues. For example Human Epididymis Protein 4 (HE4) is expressed across multiple epithelial tissues, where it is involved in innate immunity²⁸ at mucosal barriers and has been associated with decrease lung function in fibrotic lung diseases²⁹. Although direct functional relationships between HEL180 and other epithelial proteins remain unclear, their common expression in epithelial tissues suggests that they may serve similar biological roles, potentially including barrier protection and immune defense at mucosal surfaces. In contrast, downregulation of IGL c3728, an Ig light chain variant, directly indicates reduced antibody production in endotoxin-exposed workers. Importantly, adaptive immune proteins rather than innate immune proteins were collectively significant mediators of endotoxin-related lung function decline. These findings align with prior studies showing that impaired adaptive immunity, such as T-cell or B-cell deficiency, is associated with worsened lung damage and delayed recovery^{30,31}. Additional studies in human populations are needed to better understand the role of adaptive immune proteins as mediators of lung function loss.

Our study has notable strengths. First, it is the largest occupational study of textile workers with the longest follow-up period to date. Our findings help expand the understanding of endotoxin-related lung function changes to the molecular level. Our exposure assessment of endotoxin is robust. Although

we used area sampling for cotton dust, a prior validation study showed a high correlation between area and personal sampling based on job titles³². Additionally, our analytical approach addresses a complex research question using multi-omics. Using proteomics profiling of cotton and silk workers, we identified protein mediators of endotoxin-related lung function decline. Sequentially, we applied low-pass WGS to mitigate the risk of unmeasured confounding, we employed MR, which further support the influence of HEL180 and IGL c3728 on lung function change found in the OLS estimates. We applied several MR methods to minimize potential bias from horizontal pleiotropy and found similar results as the gold standard IVW.

Our study has several limitations. First, due to the nature of an occupational longitudinal study design, selection biases, such as the healthy worker survival effect and loss to follow-up, are difficult to avoid. In this follow-up study, we observed that the 413 participants who had proteomics and genotyping assessed had better lung functions and fewer respiratory symptoms than those who did not participate at baseline. Such a difference could bias the effect estimate toward the null. Second, although we implemented a robust exposure assessment, all participants were retired and had ceased occupational exposure to endotoxin by the time of our most recent survey in 2016. Proteomic measurements were only taken at a single cross-sectional time point. Our previous research suggested that some

effects of endotoxin exposure on lung damage may be partially reversible¹¹. While our focus was on the long-term consequences of exposure, by the time proteomic data were collected, some adverse health effects, whether related to lung function or protein expression, may have already partially recovered. Consequently, our estimates and findings may be conservative. Future studies could explore longitudinal changes in protein expression through measurements at two or more time points. Moreover, due to the uniqueness of our cohort, replicating our results poses a challenge, particularly with long-term repeated measurement of endotoxin. This may limit the generalizability of our results.

Despite decades of research, there is currently no consensus on occupational exposure limits for endotoxins³³. In Europe, the Dutch Expert Committee on Occupational Safety (DECOS) initially recommended a limit of 50 EU/m³ in 1998, which was raised to 200 EU/m³ in 2005 and later reduced to 90 EU/m³ in 2010³³. In the U.S., although the Occupational Safety and Health Administration has established permissible exposure limits (PELs) for cotton dust, there are no PEL standards for endotoxin exposure. In our cohort, most cotton textile workers worked in environments where endotoxin levels exceeded the 8-hour time-weighted average of 50 EU/m³ until retirement or cessation of exposure. Our results show a range of immunologic changes and respiratory effects in response to occupational exposure to airborne endotoxins. To protect the health of

workers exposed to endotoxin, further research and regulatory actions are needed to establish clear exposure limits in workplaces.

Conclusion

This study provides biological insights into mechanisms underlying endotoxin-related lung function decline. Our findings suggest chronic endotoxin exposure impairs broad immunity and that through adaptive immune proteins, endotoxin reduces lung function. We also identified two potential adaptive immune protein mediators. Our findings also emphasize the importance of workplace interventions to reduce or eliminate endotoxin exposure and promote respiratory health among susceptible populations in LMICs. However, given the caveats of our study, caution is recommended when interpreting the findings.

Declarations

Ethics approval and consent to participate

Ethical approval was obtained from the Institutional Review Boards of the Harvard School of Public Health and Shanghai Putuo District People's Hospital. All participants were provided with detailed information about the purpose, procedures, potential risks, and benefits of the study.

Consent for publication

All personal identifiers were removed to ensure participant confidentiality.

Availability of data and materials

Data used in this study are available upon reasonable request and subject to approval by the Institutional Review Board of the Harvard T.H. Chan School of Public Health. A data use agreement must be in place prior to data transfer. For inquiries, please contact Dr. David Christiani at dchris@hsph.harvard.edu. The detailed statistical analysis plan is included in the supplemental materials. Programming code used in the study is available upon request.

Competing Interest

No competing interest.

Funding

This work was funded by CDC/NIOSH grants R01OH002421 and T42OH008416, and NIH/NIEHS grant ES000002. This research was supported by the Intramural Research Program of the National Institutes of Health (NIH). The contributions of the NIH author(s) are considered Works of the United States Government. The findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services.

Author Contribution:

DCC established the cohort and supervised the project. FZ, PSL, and LS collected data by administering the surveys and conducting spirometry examinations. HW and GG conducted statistical analysis. HW, JL, PG, ZL, DCC, and JYYW contributed to the conceptual design of this manuscript and interpretation of the results. HW prepared the manuscript. All authors reviewed and approved the final version of the manuscript.

Acknowledgements

We declare no conflict of interest. We express our gratitude to all the dedicated staff, past and present, in Shanghai and Boston, whose unwavering commitment has contributed to the success of our project. Our appreciation also extends to all the participants who generously shared their time, making this research possible.

Reference

1. Grand View Research. Textile market size, share & analysis: industry research report, 2027 [Internet]. San Francisco (CA): Grand View Research; 2024 [cited 2025 Aug 4]. Available from: <https://www.grandviewresearch.com/industry-analysis/textile-market>
2. Fletcher K. Sustainable fashion and textiles: Design journeys. Earthscan; 2008.
3. Miyake K. Endotoxin recognition molecules, Toll-like receptor 4-MD-2. *Semin Immunol*. 2004;16(1):11-16. doi:10.1016/j.smim.2003.10.007
4. Buchholz BM, Billiar TR, Bauer AJ. Dominant role of the MyD88-dependent signaling pathway in mediating early endotoxin-induced murine ileus. *Am J Physiol Gastrointest Liver Physiol*. 2010;299(2):G531-G538. doi:10.1152/ajpgi.00060.2010
5. Liu T, Zhang L, Joo D, Sun SC. NF- κ B signaling in inflammation. *Signal Transduct Target Ther*. 2017;2:17023-. doi:10.1038/sigtrans.2017.23
6. Christiani DC, Eisen EA, Wegman DH, et al. Respiratory disease in cotton textile workers in the People's Republic of China. II. Pulmonary function results. *Scand J Work Environ Health*. 1986;12(1):46-50. doi:10.5271/sjweh.2175
7. Christiani DC, Ye TT, Wegman DH, Eisen EA, Dai HL, Lu PL. Cotton dust exposure, across-shift drop in FEV1, and five-year change in lung function. *Am J Respir Crit Care Med*. 1994;150(5 Pt 1):1250-1255. doi:10.1164/ajrccm.150.5.7952548
8. Christiani DC, Wang XR, Pan LD, et al. Longitudinal changes in pulmonary function and respiratory symptoms in cotton textile workers. A 15-yr follow-up study. *Am J Respir Crit Care Med*. 2001;163(4):847-853. doi:10.1164/ajrccm.163.4.2006063
9. Wang XR, Eisen EA, Zhang HX, et al. Respiratory symptoms and cotton dust exposure; results of a 15 year follow up observation. *Occup Environ Med*. 2003;60(12):935-941. doi:10.1136/oem.60.12.935
10. Wang XR, Zhang HX, Sun BX, et al. A 20-year follow-up study on chronic respiratory effects of exposure to cotton dust. *Eur Respir J*. 2005;26(5):881-886. doi:10.1183/09031936.05.00125604
11. Lai PS, Hang JQ, Valeri L, et al. Endotoxin and gender modify lung function recovery after occupational organic dust exposure: a 30-year study. *Occup Environ Med*. 2015;72(8):546-552. doi:10.1136/oemed-2014-102579
12. Anyfantis ID, Rachiotis G, Hadjichristodoulou C, Gourgoulialis KI. Respiratory Symptoms and Lung Function among Greek Cotton Industry Workers: A Cross-Sectional Study. *Int J Occup Environ Med*. 2017;8(1):32-38. doi:10.15171/ijoem.2017.888
13. Nafees AA, Allana A, Kadir MM, Potts J, Minelli C, Semple S, De Matteis S, Burney P, Cullinan P. A cluster randomised controlled trial to reduce respiratory effects of cotton dust exposure among textile workers: the MultiTex RCT study. *Eur Respir J*. 2024 Jan 18;63(1):2301028. doi: 10.1183/13993003.01028-2023. PMID: 37857425.
14. Smid T, Heederik D, Houba R, Quanjer PH. Dust- and endotoxin-related acute lung function changes and work-related symptoms in workers in the animal feed industry. *Am J Ind Med*. 1994;25(6):877-888. doi:10.1002/ajim.4700250611

15. Zhao M, Wei L, Zhang L, et al. Proteomic biomarkers of long-term lung function decline in textile workers: a 35-year longitudinal study. *J Expo Sci Environ Epidemiol.* 2025;35(4):602-610. doi:10.1038/s41370-024-00721-7
16. Graham BL, Steenbruggen I, Miller MR, et al. Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. *Am J Respir Crit Care Med.* 2019;200(8):e70-e88. doi:10.1164/rccm.201908-1590ST
17. The Gene Ontology Consortium. The Gene Ontology knowledgebase in 2023. *Genetics.* 2023 May 4;224(1):iyad031. DOI: [10.1093/genetics/iyad031](https://doi.org/10.1093/genetics/iyad031)
18. Rubinacci S, Hofmeister RJ, Sousa da Mota B, Delaneau O. Imputation of low-coverage sequencing data from 150,119 UK Biobank genomes. *Nat Genet.* 2023;55(7):1088-1090. doi:10.1038/s41588-023-01438-3
19. Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 2023;51(D1):D638-D646. doi:10.1093/nar/gkac1000
20. Prim RC. Shortest connection networks and some generalizations. *Bell Syst Tech J.* 1957;36(6):1389-401.
21. Massara GP, Di Matteo T, Aste T. Network filtering for big data: Triangulated maximally filtered graph. *J Complex Netw.* 2016;5(2):161-78.
22. Dai JY, Stanford JL, LeBlanc M. A multiple-testing procedure for high-dimensional mediation hypotheses. *J Am Stat Assoc.* 2022;117(537):198-213. doi:10.1080/01621459.2020.1765785
23. Liu Z, Shen J, Barfield R, Schwartz J, Baccarelli AA, Lin X. Large-Scale Hypothesis Testing for Causal Mediation Effects with Applications in Genome-wide Epigenetic Studies. *J Am Stat Assoc.* 2022;117(537):67-81. doi:10.1080/01621459.2021.1914634
24. Yang H, Liu Z, Wang R, Lai E-Y, Schwartz J, Baccarelli AA, Huang YT, Lin X. Causal mediation analysis for integrating exposure, genomic, and phenotype data. *Annu Rev Stat Appl.* 2024;12:337-360. doi:10.1146/annurev.statistics.040622.031653
25. Szafran B, Borazjani A, Lee JH, Ross MK, Kaplan BL. Lipopolysaccharide suppresses carboxylesterase 2g activity and 2-arachidonoylglycerol hydrolysis: A possible mechanism to regulate inflammation. *Prostaglandins Other Lipid Mediat.* 2015;121(Pt B):199-206. doi:10.1016/j.prostaglandins.2015.09.005
26. Lai PS, Fresco JM, Pinilla MA, et al. Chronic endotoxin exposure produces airflow obstruction and lung dendritic cell expansion. *Am J Respir Cell Mol Biol.* 2012;47(2):209-217. doi:10.1165/rcmb.2011-0447OC
27. Janeway CA Jr, Travers P, Walport M, Shlomchik MJ. Immunobiology: The immune system in health and disease. 5th ed. New York: Garland Science; 2001.
28. Bingle, L., Cross, S. S., High, A. S., Wallace, W. A., Rassl, D., Yuan, G., Hellstrom, I., Campos, M. A., & Bingle, C. D. (2006). WFDC2 (HE4): a potential role in the innate immunity of the oral cavity and respiratory tract and the development of adenocarcinomas of the lung. *Respiratory research*, 7(1), 61. <https://doi.org/10.1186/1465-9921-7-61>
29. Zhang, M., Zhang, L., E, L., Xu, K., Wang, X. F., Zhang, B., Su, J., & Meng, Z. (2020). Increased levels of HE4 (WFDC2) in systemic sclerosis: a novel biomarker reflecting interstitial lung disease severity?. *Therapeutic*

- advances in chronic disease*, 11, 2040622320956420.
<https://doi.org/10.1177/2040622320956420>
30. Aziz M, Ode Y, Zhou M, et al. B-1a cells protect mice from sepsis-induced acute lung injury. *Mol Med*. 2018;24(1):26. Published 2018 May 29. doi:10.1186/s10020-018-0029-2
 31. Ovsyannikova IG, Pankratz VS, Salk HM, Kennedy RB, Poland GA. HLA alleles associated with the adaptive immune response to smallpox vaccine: a replication study. *Hum Genet*. 2014;133(9):1083-1092. doi:10.1007/s00439-014-1449-x
 32. Mehta AJ, Wang XR, Eisen EA, et al. Work area measurements as predictors of personal exposure to endotoxin and cotton dust in the cotton textile industry. *Ann Occup Hyg*. 2008;52(1):45-54. doi:10.1093/annhyg/mem061
 33. Farokhi A, Heederik D, Smit LAM. Respiratory health effects of exposure to low levels of airborne endotoxin - a systematic review. *Environ Health*. 2018;17(1):14. Published 2018 Feb 8. doi:10.1186/s12940-018-0360-7
 34. Dutch Expert Committee on Occupational Safety (DECOS), Health Council of the Netherlands. Health-based recommended occupational exposure limit for endotoxins. The Hague: DECOS; 2010 Jul 15. Report No.: 2010/04OSH.

Table 1: Characteristics of participants with proteomics assessed in the Shanghai Textile Worker Study (N=413)

	Cotton N=221	Silk N=192
Sex		
Male	73 (33 %)	57 (30 %)
Female	148 (67 %)	135 (70 %)
Age in 2016 (year), Mean (SD)	70 ($\pm$ 8.7)	70 ($\pm$ 8.6)
Cumulative endotoxin concentration (1000 EU/m³-year), Mean (SD)	46 ($\pm$ 40)	0 ($\pm$ 0)
Smoking status in 2016		
Former	32 (14 %)	24 (12 %)
Current	34 (15 %)	23 (12 %)
Never	155 (70 %)	145 (76 %)
Total smoking pack-year in 2016 (pack-year), Mean (SD)	9.9 ($\pm$ 22)	6.6 ($\pm$ 15)

FEV-1 in 2011 (ml), Mean (SD)	2328 ($\pm$ 619)	2319 ($\pm$ 571)
FVC in 2011 (ml), Mean (SD)	3023 ($\pm$ 774)	3002($\pm$ 741)
Percent predicted* FEV-1 in 2011 (%),Mean (SD)	94.8 ($\pm$ 16.2)	97.8 ($\pm$ 15.7)
Percent Predicted FVC in 2011 (%), Mean (SD)	82.6 ($\pm$ 13.2)	85.0 ($\pm$ 12.5)
FEV-1 in 2016 (ml), Mean (SD)	2100 ($\pm$ 590)	2100 ($\pm$ 550)
FVC -1 2016 (ml), Mean (SD)	2757 ($\pm$ 738)	2772 ($\pm$ 722)
Percent predicted FEV-1 in 2016 (%), Mean (SD)	91.3 ($\pm$ 16.7)	93.8 ($\pm$ 15.6)
Percent predicted FVC in 2016 (%), Mean (SD)	74.9 ($\pm$ 13.2)	78.1 ($\pm$ 12.0)
Annual change in FEV-1 between 2011 and 2016 (ml/year), Mean (SD)	-25 ($\pm$ 11)	-23 ($\pm$ 9.6)
Annual change in FVC between 2011 and 2016 (ml/year), Mean (SD)	-42.9 ($\pm$ 41)	-36 ($\pm$ 43)

Definition of abbreviations: EU = endotoxin unit; FEV-1= forced expiratory volume in 1 second; SD=standard deviation. FVC= Force Vital Capacity

Data are expressed as mean ($\pm$ SD) for continuous variables and n (%) for categorical variables.*Percent predicted FEV-1 and FVC were calculated from Global Lung Initiative race-neutral formula.

Table 2. Protein mediators of endotoxin-related lung function changes and summary of mediation analysis results

Protein Names	IGL c3728_light_IGKV4-1_IGKJ1	Epididymis luminal protein 180
Protein ID	A0A5C2G374	B6EDE2
P-value of HDMT	9.41×10^{-6}	6.83×10^{-6}
P-value of Modified DACT	1.42×10^{-5}	8.71×10^{-6}
Average causal mediation effect (95% CI)	-0.73 (-1.29, -0.29)	-0.717 (-1.21, -0.31)
Average direct effect (95% CI)	-0.82 (-2.55, 0.69)	-0.827 (-2.63, 0.69)
Total effect (95% CI)	-1.54 (-3.31, -0.03)	-1.544 (-3.27, -0.06)
Proportion mediated (p-value)	46.9% (0.05)	46.4% (0.04)

Definition of abbreviations: CI=Confidence Interval; DACT= divide-aggregate composite-null test; HDMT = high dimensional mediation analysis.

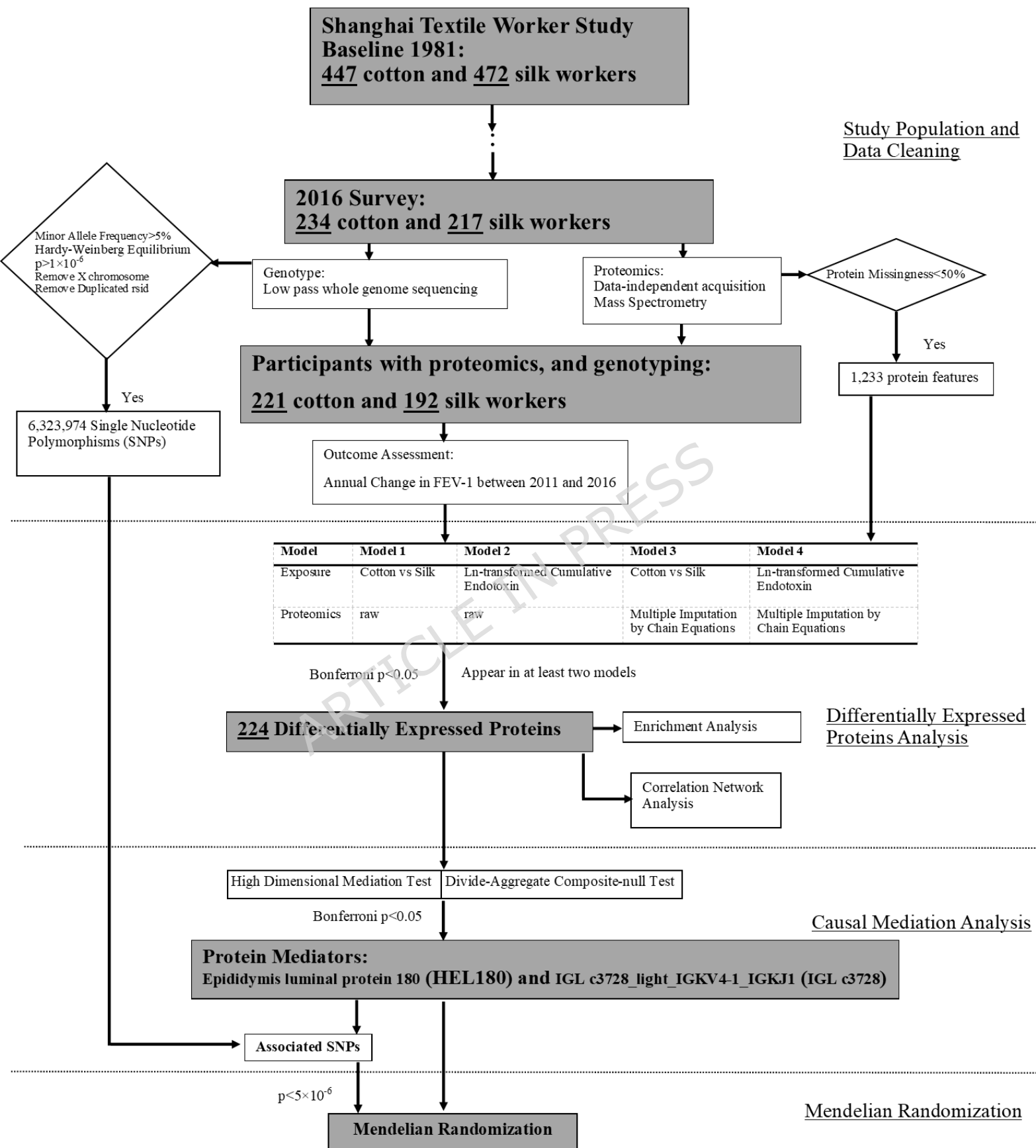
ARTICLE IN PRESS

Figure Legends:

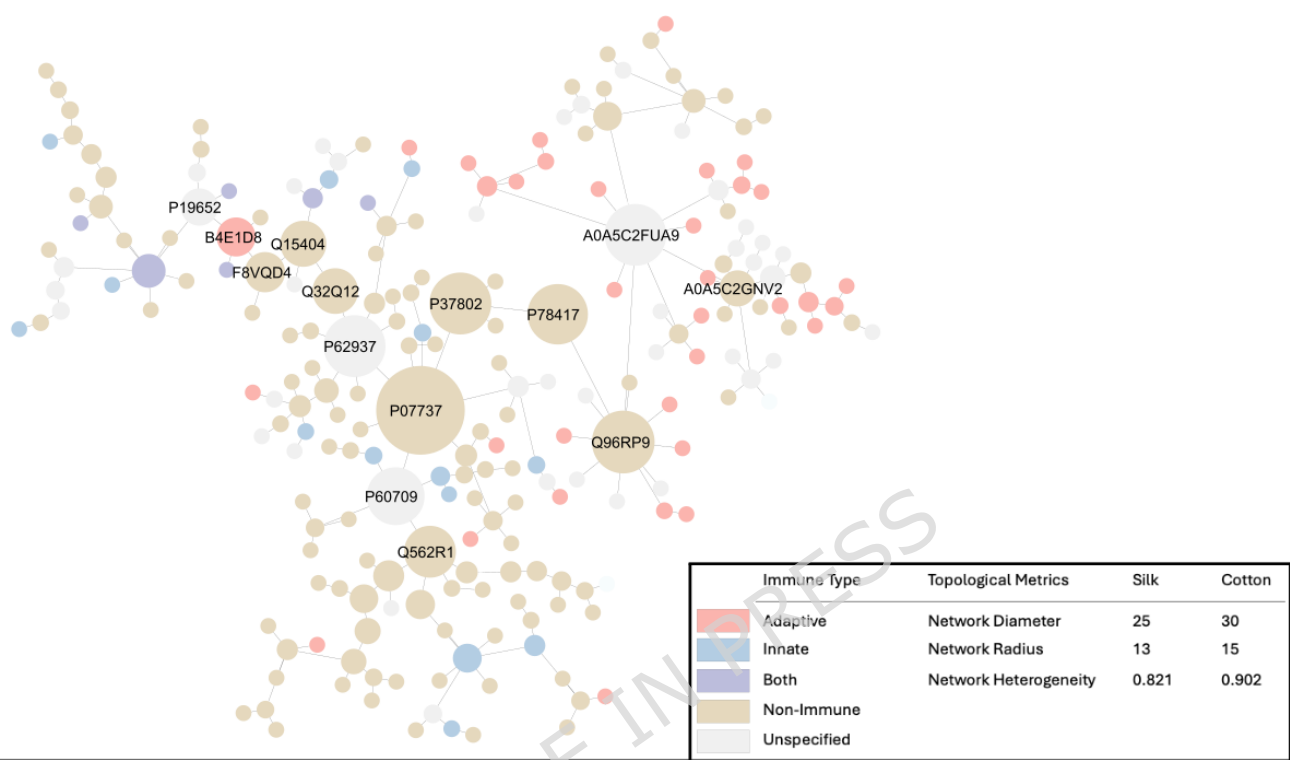
Figure 1. Overall study design and flow chart of statistical analysis.

Figure 2. Protein-protein correlation network of differentially expressed proteins associated with occupation endotoxin exposure, stratified by occupation. The network was constructed using Prim's maximum spanning tree. Only the top 15 proteins by betweenness centrality are labeled.

Figure 3. Forest plot of overall estimates for the effect of a two-fold increase in protein mediator expression on lung function change, based on Mendelian Randomization (MR) analyses and adjusted ordinary least squares regression. MR methods include robust adjusted profile score, MR-PRESSO, MR-Lasso, MR-Egger, contamination mixture, and inverse-variance weighted approaches.



A. Silk



B. Cotton

