Exposome, 2024, **4(1),** osae002
https://doi.org/10.1093/exposome/osae002
Advance Access Publication Date: 12 February 2024
Research Article

Questionnaire-based exposome-wide association studies for common diseases in the Personalized Environment and Genes Study

Dillon Lloyd[1]; John S. House [1]; Farida S. Akhtari[1]; Charles P. Schmitt [2]; David C. Fargo[3]; Elizabeth H. Scholl[4]; Jason Phillips[4]; Shail Choksi[4]; Ruchir Shah[4]; Janet E. Hall[5,†]; Alison A. Motsinger-Reif [1,†,*]

[1]Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, Durham, NC, USA
[2]Division of Translational Toxicology, National Institute of Environmental Health Sciences, Durham, NC, USA
[3]Office of the Director, National Institute of Environmental Health Sciences, Durham, NC, USA
[4]Sciome LLC, Durham, NC, USA
[5]Clinical Research Branch, National Institute of Environmental Health Sciences, Durham, NC, USA

*To whom correspondence should be addressed: Email: alison.motsinger-reif@nih.gov
†Equal contribution.

Abstract

The exposome collectively refers to all exposures, beginning *in utero* and continuing throughout life, and comprises not only standard environmental exposures such as point source pollution and ozone levels but also exposures from diet, medication, lifestyle factors, stress, and occupation. The exposome interacts with individual genetic and epigenetic characteristics to affect human health and disease, but large-scale studies that characterize the exposome and its relationships with human disease are limited. To address this gap, we used extensive questionnaire data from the diverse North Carolina-based Personalized Environment and Genes Study (PEGS, $n = 9,429$) to evaluate exposure associations in relation to common diseases. We performed an exposome-wide association study (ExWAS) to examine single exposure models and their associations with 11 common complex diseases, namely allergic rhinitis, asthma, bone loss, fibroids, high cholesterol, hypertension, iron-deficient anemia, ovarian cysts, lower GI polyps, migraines, and type 2 diabetes. Across diseases, we found associations with lifestyle factors and socioeconomic status as well as asbestos, various dust types, biohazardous material, and textile-related exposures. We also found disease-specific associations such as fishing with lead weights and migraines. To differentiate between a replicated result and a novel finding, we used an AI-based literature search and database tool that allowed us to examine the current literature. We found both replicated findings, especially for lifestyle factors such as sleep and smoking across diseases, and novel findings, especially for occupational exposures and multiple diseases.

Keywords: exposome-wide association study (ExWAS); Personalized Environment and Genes Study (PEGS); occupational exposures; exposome data; common diseases; environmental exposures.

Introduction

The exposome, which represents all exposures over the course of an individual's life, has substantial effects on human health and disease.[1,2] Although the effects of general lifestyle factors such as diet and exercise[3-5] and aggregated exposures such as air[6-8] and water pollution[9] have been well-studied, occupational exposures, as well as exposures from hobby activities, have not been thoroughly examined on a large scale. While prior work has used data from exposome surveys on multiple exposures and phenotypes to assess individual exposures,[10-13] cohorts with comprehensive exposome data are rare, and few studies have examined exposome associations with multiple complex diseases. Existing studies have mainly focused on single traits such as asthma[14] or other biological phenotypes.[15-17] Further, despite recent emphasis on the importance of representing multi-ancestry and varied socioeconomic populations[18-21] in exposure science, study populations with a large number of underrepresented individuals remain rare. To address these gaps, we conducted exposome-wide association studies (ExWAS) to examine associations between the exposome and common, complex diseases in the Personalized Environment and Genes Study (PEGS),[10,11,22-24] a unique North Carolina-based multi-ancestry cohort with extensive health and exposure data.

An ExWAS is analogous to a genome-wide association study (GWAS) but considers trait associations with exposure variation instead of genetic variation.[25] In a typical ExWAS analysis, for a disease of interest, data for single exposures are modeled separately using an appropriate regression model, and then procedures such as false discovery rate (FDR) correction are used to control for multiple testing.[26] In data from the PEGS cohort, we selected 11 common, complex human diseases with sufficient prevalence (and thus statistical power) to enable discovery. Using questionnaire-based data on lifestyle factors and exposures, namely sleep issues, smoking, socioeconomic factors, mood- and fatigue-related factors, occupational and hobby exposures, and infectious diseases, we conducted ExWAS for each phenotype. This is analogous to phenome-wide association studies (PheWAS), which invert the concept of a GWAS by searching for phenotypes associated with

Received: November 20, 2023. **Accepted:** January 1, 2024
Published by Oxford University Press 2024.

specific genetic variants across the range of human phenotypes, or the phenome.[27-29] Motivated by the same goal, our ExWAS also assess associations of important exposures across phenotypes.

Exploring the links between environmental exposures and multiple diseases is crucial for public health and scientific research.[30] This allows policymakers and public health agencies to focus on shared risk factors rather than factors specific to individual diseases, improving resource allocation. Identifying common environmental exposures also helps develop targeted prevention strategies. Reducing exposure to shared risk factors can significantly reduce the overall disease burden by simultaneously targeting multiple diseases. Uncovering these links also supports better public health messaging[31] to educate and inform the public about these shared risk factors. This enables individuals to take proactive measures and make informed lifestyle choices.

Additionally, understanding how exposures are associated with multiple diseases helps researchers prioritize research on common environmental exposures. Such across-phenotype interrogation can uncover the physiological effects of shared risk factors, with the results potentially used to develop new therapeutic targets or inform intervention strategies with broad applications. Additionally, identifying common environmental exposures associated with multiple diseases aids in early detection and treatment as these exposures are often risk factors throughout life. Knowledge of shared risk factors allows healthcare providers to closely monitor at-risk individuals and intervene quickly to prevent or manage disease progression. Finally, understanding how environmental exposures affect biological systems and tissues can help personalize treatment and eventually advance precision medicine with therapeutic interventions tailored to an individual's exposure history and genetic makeup.

Here, we present ExWAS results across multiple, high-prevalence diseases. We identify individual factors for each disease and conduct multi-exposure modeling to identify the most important mixtures associated with each disease. We also contextualize our results across diseases within the current literature. Given the high dimensionality of not only our data but also our results, we conducted a comprehensive literature search using artificial intelligence and natural language processing tools to empower this contextualization. Taken together, the results demonstrate the value of performing ExWAS across highly prevalent diseases in the PEGS cohort to interrogate both new and established relationships between the exposome and disease and sharing these results for the larger community to examine and use for further study.

Methods

Materials

Study participants

We used self-reported exposure and phenotype information from 9429 PEGS participants (data freeze 2). PEGS is a North Carolina-based cohort comprising demographically representative, diverse participants in terms of age, education, socioeconomic status, race, and sex. Table 1A displays the components collected by each survey, and Table 1B provides participant demographics of the sample used for the study.

Exposome data

PEGS collects exposome data using three surveys. The Health and Exposure Survey, administered beginning in 2013, includes questions on general health, family medical history, lifestyle factors, and occupational exposures. Beginning in 2017, two additional surveys were administered. The Internal Exposome Survey collects information on medication and additional lifestyle factors such as diet and sleep patterns, and the External Exposome Survey collects information on chemical and environmental exposures. There is a substantial overlap of participants who have completed both the Internal and External Exposome Surveys as participants were chosen for subsequent surveys from those who completed the Health and Exposure Survey. Figure 1 shows the type of information collected by each survey, and Table 2 summarizes the exposome information collected by the PEGS surveys. The full questionnaires are available at https://www.niehs.nih.gov/research/atniehs/labs/crb/studies/pegs/about/data/index.cfm.

Phenotype definitions

In the current study, we defined cases and controls based on survey responses for 11 common, complex human traits, namely allergic rhinitis, asthma, bone loss, fibroids, iron-deficient anemia, lower GI polyps, migraines, ovarian cysts, type 2 diabetes, high cholesterol, and hypertension. We selected these traits based on power calculations that were used to define a prevalence cut-off for inclusion in the study. We used SAS proc(power) with a logistic regression model, assuming a Bonferroni corrected alpha based on the number of exposures tested in the ExWAS. We assumed a 10% environmental exposure frequency and 10% trait prevalence. These cut-off values resulted in 80% power to detect an odds ratio greater than 1.375, which is a reasonable effect size given previous ExWAS conducted in the PEGS data.[11] We defined cases (response = "Yes") and controls (response = "No") for each phenotype using the Health and Exposure Survey question that asked whether a participant has ever been diagnosed by a doctor or physician with the condition. We determined the inclusion/exclusion criteria for each phenotype from the results of a literature review. Supplementary Figures S1-S11 outline the criteria used to define cases and controls for each phenotype.

Statistical analysis

Exposome-wide association studies (ExWAS)

In quality control conducted prior to analysis, we filtered out questions unrelated to health or exposures, redundant questions, and questions with free-text answers. We also excluded questions that were nested and dependent on the answer to a previous question due to insufficient sample sizes for individual response levels. We further excluded from the analysis questions with fewer than 30 responses or invariant responses. We also excluded exposures with fewer than five cases for either the exposure or a factor level of the exposure. After this filtering, 281 variables remained for the Health and Exposure Survey, 359 variables for the Internal Exposome Survey, and 438 variables for the External Exposome Survey. We calculated correlations between all variable pairs within and across surveys.

We conducted ExWAS to test the association of individual variables with each phenotype. We performed logistic regression for each phenotype and exposure with the following equation, adjusting for age, income, self-reported race/ethnicity, and sex, and applied a Benjamini-Hochberg FDR of 0.10 to define significance.

$$Trait \sim Exposure + Age + Sex + Income + Race + \in$$

Deletion/substitution/addition (DSA) algorithm for multi-exposure models

Individual exposures never occur in isolation, so we performed variable selection to identify exposure mixtures that are strongly associated with the 11 diseases. To identify concurrent associations of

Table 1. Personalized environment and genes study (PEGS) data components by survey and participant demographic information

A

Data component	Description	Number of participants
Health and Exposure Survey	Demographics, individual and family health history, environmental exposures, socioeconomic status, and lifestyle factors	9426
External Exposome Survey	Residential and occupational environmental exposures	3579
Internal Exposome Survey	Medication use, physical activity, stress, sleep, diet, genetics, and reproductive history	3034

B

Demographic variable	Health and exposure	External exposome	Internal exposome
Number of participants	9426	3579	3034
Sex, *N* (%)			
Female	6317 (67)	2250 (70)	1927 (70)
Male	3103 (33)	986 (30)	843 (30)
Income, *N* (%)			
Over 80 000 per year	2747 (30)		
Under 80 000 per year	6348 (70)		
Race, *N* (%)			
White	6711 (73)	2703 (84)	2334 (85)
Black	2055 (22)	338 (11)	272 (10)
Other	412 (5)	161 (5)	136 (5)
Ethnicity, *N* (%)			
Non-Hispanic/Non-Latino	8881 (98)	3435 (97)	2920 (93)
Hispanic/Latino	371 (2)	118 (3)	93 (7)
Age (years), mean (SD)	50.17 (15.96)	50.23 (14.34)	50.28 (14.34)
Body mass index (BMI), mean (SD)	28.65 (6.82)	28.02 (6.70)	27.91 (6.67)

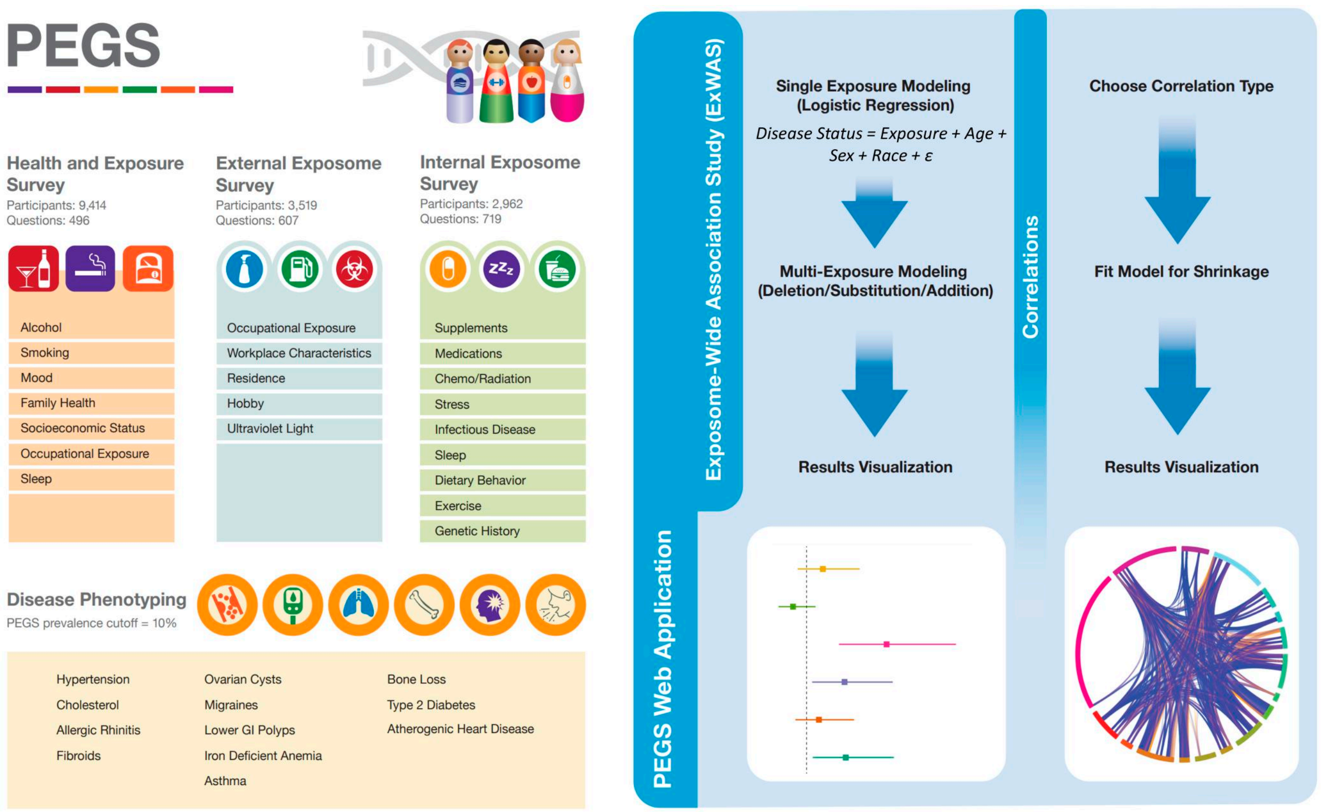

Figure 1. Overview of the PEGS cohort and PEGS Explorer. PEGS collects data on participants' internal and external exposures and health. We performed phenotyping for diseases that met a prevalence cutoff in the overall cohort. PEGS Explorer displays results for exposome-wide association studies (ExWAS) and correlation analyses. Users can view the results through an interactive website.

exposures with diseases, we conducted multi-exposure modeling using all ExWAS-significant exposures (FDR ≤ 0.10) as input to a deletion/substitution/addition (DSA) algorithm.[32] Through a series of deletions, substitutions, and additions to a logistic regression model, a DSA algorithm returns a list of ExWAS-significant variables that, in combination, are the simplest and most predictive

Table 2. Exposome information collected by each Personalized Environment and Genes Study (PEGS) survey

External Exposome Survey	
Characteristics of current residence	
	Central heat, fuel type: natural gas, gas, electricity, oil or kerosene, wood, solar energy, other
	fireplace or wood-burning stove, fuel type: electricity, natural gas, propane
Cooling	Air conditioning
Water	Source for drinking water, water filtering system, source for bathing/showering
Water and dampness	Water or dampness (past 12 months), mildew odor or musty smell, repair or renovation, mold pollution, frequency of HEPA filter and vacuum use
Walls and flooring	Any painting, wallpaper or flooring repair (past 12 months), carpet, plastic floor, textile wall material, plastic wall material
Garage and basement	Garage, basement, crawlspace, gas-powered devices
Pets	Indoor pets
Pesticides and insecticides	Regular use of insecticides or pesticides
Surrounding area	Traffic intensity, distance from bus station, airport, dry cleaners, farm, gas station, golf course, greenhouse, hazard waste, power lines, incinerator, landfill, tannery, military base, oil refinery, paper mill, poultry plant, coal plant, nuclear plant, or factory
Agricultural property use	Property used as a farm or orchard before/while living there
Chemical and metal exposures at work	
Solvent	Benzene, chloroform, chloroprene, dichlorobenzene, ethyl benzene, ethyl dichloride, perchloroethylene, toluene, trichloroethylene, xylenes
Lubricating oil	Brake fluid, transmission fluid, hydraulic fluid, motor oil, other lubricating oil
Cleaning liquid	Bleach, ammonia, carbon tetrachloride, other
Heavy metal	Arsenic, beryllium cadmium chromates, lead, mercury, nickel, other
Alcohols	Isopropanol, methanol, ethanol, butanol, other
Pesticides/fumigant	Ethyl dibromide, insecticides, fungicides, herbicides, ethyl fumigants, rodenticides, other
Plastic production	BPA, vinyl chloride, styrene, phosgene, phenol, toluene diisocyanate, methylene, other
Dust	Coal, fiberglass, rock dust, silica, talc, other
Emission	Nitrous oxide, carbon dioxide, carbon monoxide, ozone, other
Carcinogens	Polybrominated biphenyls (PBBs), polychlorinated biphenyls, radiations, x-rays, welding fumes, other
Acids	Hydrochloric, sulfuric, phosphoric, acetic, nitric, other
Alkalis	Sodium hydroxide, calcium hydroxide, potassium hydroxide, magnesium hydroxide, other
Stains and varnishes	Shellac, wood stain, resin varnish, polyurethane, lacquer, acrylic, other
Paints and paint thinners	Primer, enamel, oil-based, acrylic, luminescent, acetone, turpentine, naphtha, methyl ethyl ketone, other
Anesthetic	Desflurane, halothane, isoflurane, nitrous, sevoflurane, other
Glues and adhesive	White glue, rubber cement, neoprene, ethylene vinyl acetate, epoxy, urethane, polyimides, cyanoacrylates, wallpaper paste, other
Soldering	Eutectic, tin zinc, lead silver, cadmium silver, flux, solder paste, solder wire, rosin, solder fumes, other
Dyes and ink	Dye: hair, leather, textile, paper
	Ink: india, inkjet, gel, pen, soy, pharmaceutical, toner, other
Use of protective gear	Gloves, mask, clothing
Workplace characteristics	
Floor material	Concrete, wood, cork, vinyl, wall-to-wall carpet, other
Wall material	Textile, plastic
Hobby related	
	Glue, solder, photo, painting, wood, ceramics, leather, fishing, mechanic, yard work, other
Ultraviolet light	
	Reaction without sunscreen, sunburn experience, skin rash, use of tan booth, use of sunlamp, weekday summer, weekend summer, # of months/year of tanning, leisure time hat, sunglass, sunscreen
Internal Exposome Survey	
Current use of supplements	
Vitamins	Multi-vitamin, A, B3, B6, B12, B complex, C, D, E,
Minerals	Calcium, chromium, iron, potassium, selenium, zinc
Supplements	Black cohosh, coenzyme Q10, fish oil, flaxseed oil, folic acid, ginkgo biloba, ginseng, glucosamine/chondroitin, melatonin, milk thistle, omega-3 fatty acids, probiotics, red yeast rice, resveratrol, St John's wort
Current use of prescription medication	
	Diabetes, thyroid disease, anxiety, depression, asthma, chronic bronchitis or emphysema (COPD), acid reflux, chronic pain, insomnia or sleep disorders, birth control, hormone replacement therapy
Chemotherapy/radiation therapy	
	Ever diagnosed with cancer
Physical activity	
	Strenuous exercise, moderate exercise, mild exercise
Stress	
	Being upset, unable to control important things, felt nervous/stressed, felt confident handling personal problems, felt that things were going your way, felt able to cope, able to control irritation, felt that you were on top of things, been angered, felt overwhelmed

(continued)

Table 2. (continued)

Internal Exposome Survey
Infectious disease
Chicken pox, chlamydia, cold sores, cryptosporidiosis, dysentery, flu, food poisoning (E. coli enteritis), food poisoning (norovirus), genital herpes, genital warts, German measles, gonorrhea, hepatitis A/B/C, HIV/AIDS, legionellosis, Lyme disease, malaria, measles, meningitis (bacterial and viral), mononucleosis, mumps, pneumonia (bacterial and virus), Rocky Mountain spotted fever, salmonella infection, shingles, staph infection, streptococcal invasive disease, syphilis, tuberculosis, ulcers, whooping cough
Sleep
Days/week napping at least 20min, waking up in the middle of the night, nighttime bathroom use, difficulty breathing, coughing, snoring, felt too cold or hot, had bad dreams, felt pain, felt stressed or anxious, affected by spouse's or significant other's sleeping habits, sleep quality, sleep medication use, trouble staying awake, stopping breathing during sleep, diagnosis of sleep disorder
Dietary behavior
Eating at or bringing food from fast food, sit-down restaurant, buffet, takeout, grocery stores, cafeterias, vending machines, on-street vendors, other
Dietary intake
Milk, cream, non-dairy coffee creamer, frozen yogurt, regular ice cream, spreads added to food or bread, yogurt, cheese, apple, avocado, banana, blueberries, cantaloupe or honeydew melon, grapes, grapefruit, orange, peaches/plums/apricots, pear, raspberries, strawberries, tomatoes, tomato or V-8 juice, tomato sauce, watermelon, beans/lentils, bell peppers, broccoli, brussels sprouts, cabbage, carrots, cauliflower, corn, eggplant/zucchini, kale, onions, peas, spinach cooked, spinach raw, string beans, tofu, winter squash, yam, eggs, beef/pork hot dogs, chicken/turkey hot dogs/sausage, chicken/turkey, bacon, salami/processed deli/sandwich meat, sausage/kielbasa/other processed meats, hamburger/ground beef, beef/pork/lamb, ham, canned tuna fish, shrimp/lobster/scallops, dark meat fish (tuna steak, mackerel, salmon, sardines, bluefish, swordfish), other fish (cod, haddock, halibut), cereals, crackers, breads, muffins, pancakes, brown rice, white rice, pasta, tortillas, French fries, potatoes (baked, boiled, or mashed), pizza, carbonated beverages, other beverages, milk chocolate, dark chocolate, candy, cookies, cake, jams/jellies/syrup/honey, peanut butter, potato chips, pretzels, peanuts, walnuts, other nuts, olive oil added to food or bread, mayonnaise, salad dressing, table sugar
Twin/triplet siblings and birth order
Twin, triplet, birth order
Genetic history
Blood type, color blind, cystic fibrosis, Down syndrome, hemochromatosis, hemophilia, hyperlipidemia, Huntington's disease, blood clot, Klinefelter syndrome, muscular dystrophy, Niemann-Pick disease, phenylketonuria, polycystic kidney disease, sickle cell, Tay-Sachs disease, thalassemia, Turner syndrome
Health and Exposure Survey
Exposures (work or daily life)
Asbestos, biohazard materials (blood, tissue, or other bodily fluids), chemicals/acid/solvent, coal or stone dust, coal tar/pitch/asphalt, diesel engine exhaust, dyes, formaldehyde, gasoline exhaust, heavy metals (lead, mercury, cadmium, arsenic), pesticides/herbicides, sand or silica, other dusty conditions, textile fibers/dust, wood dust, x-ray/radioactive materials
Alcohol and smoking
Smoked at least 100 cigarettes in a lifetime, tobacco, cigars, indoor exposure, number of smokers at home, had at least one drink of any kind of alcohol in a lifetime
Sleep
Number of nights trouble sleeping/week
Socioeconomic status
Highest education level for self and parents
Mental health/mood
Mental health history (self and family members), ever experienced unusual hyperactivity, irritability, confidence, sleepiness, talkative, having unusual thoughts, being distracted, having energy, being active, having sex, engaging in risky behavior, spending issues, social problems

multi-exposure model for a phenotype of interest. The method employs random cross-validation to find a model that is both parsimonious and predictive of the trait of interest. The use of cross-validation and the combination of forward and backward selection allows for a comprehensive look at the exposure space. Using the DSA algorithm, we conducted step-wise regression with 10-fold cross-validation to determine which exposures were associated with each disease of interest in the simplest model. We implemented the algorithm using the DSA package in R.

Causaly

We used Causaly, an artificial intelligence software and database, to gather evidence for contextualizing the ExWAS results. The high dimensionality of both the data and the significant results poses a challenge for fully comprehending and contextualizing the results. Causaly allows users to search the literature using specific terms and classifies the associations between exposures and the disease of interest using advanced artificial intelligence and natural language processing. We limited our search to the clinical term for each phenotype of interest (eg, osteopenia for bone loss) to exclude more narrowly defined conditions. For example, searches using the term "uterine fibroids" excluded results related to cancerous tumors. This ensured the search terms closely matched the phenotype definitions. For Causaly searches related to exposures, we took a contrasting approach and used the general, inclusive categories for exposures. This ensured the search terms captured all research on the exposures associated with specific diseases in the ExWAS results and remained pertinent to the studied phenotypes. For example, PEGS asks about exposure to heavy metals as a general question, so our Causaly query included the generic term "heavy metals" as well as terms for specific heavy metals such as lead and cadmium.

Causaly characterizes the current literature by associations with multiple disease outcomes. The results include summaries of the direction of association effects and shared terms across diseases and exposures. This allowed us to contextualize our results rapidly and quantitatively to better examine exposomic relationships. We used the results of our search queries to identify whether associations were new to the literature or already established. We further contextualized the results based on whether an association was identified in between one and 10 previous studies or more than 10 studies. The specific search terms used for each phenotype can be found in Supplementary Table S1. We accessed the Causaly database in October 2023.

Data availability

The data underlying this article are available to researchers engaging in collaborative projects with PEGS. Information about submitting proposals for collaborative research is available at: https://www.niehs.nih.gov/research/clinical/studies/pegs/collaboration/proposal/index.cfm. We concurrently developed a web-based tool to further interrogate the ExWAS results with the present study. Results from the ExWAS conducted as part of this work can be explored with the PEGS Explorer web application that can be accessed at: https://www.niehs.nih.gov/research/clinical/studies/pegs/index.cfm under the About PEGS tab. Details on PEGS Explorer can be found in Lloyd *et al.* in this issue. The Causaly searches can be accessed with a valid license for the software.

Results

ExWAS results

The significant results from the ExWAS analysis included associations of each of the 11 considered diseases and exposures across surveys and in topical categories ranging from occupational exposures to lifestyle and socioeconomic factors. Figure 2 shows the distribution of *P* values and the effect direction for the results across phenotypes. There was substantial overlap across phenotypes in occupational exposures and socioeconomic, lifestyle, and fatigue- and mood-related factors. UpSet plots for each survey that show the overlap between exposures and diseases are shown in Figure 3. Supplemental file S1 shows the results for each survey in an Excel worksheet.

In the discussion, we highlight and contextualize the results that remained significant after FDR correction. We highlight expected results that replicate well-established associations in the literature and novel associations that have not been previously reported. Because of their high dimensionality, the ExWAS results cannot be fully described in this manuscript. The PEGS Explorer web application enables exploration of the full results

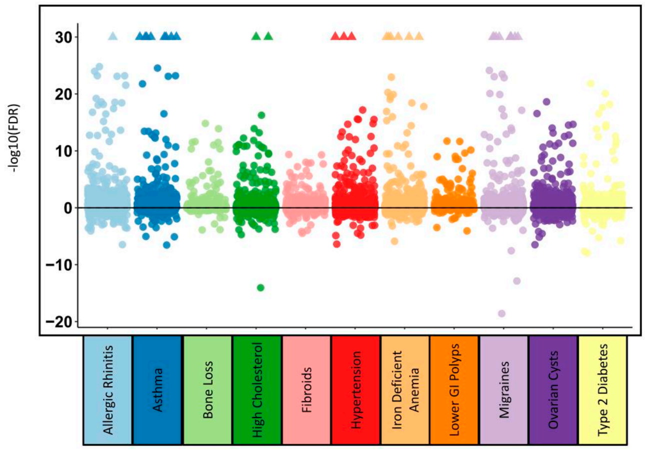

Figure 2. Miami plot for all surveys and traits. The y-axis represents –log10 (FDR-adjusted *P* values) with the direction of effect indicating association or inverse association with the trait of interest. Points represented with triangles exceeded –axis limits for a particular exposure/phenotype association.

by providing access to the comprehensive ExWAS results for disease-phenotype associations and correlations among individual exposures. Many well-established lifestyle factors related to diet, sleep, and exercise were significantly associated with diseases in the ExWAS results, providing assurance that the ExWAS have sufficient power to detect established associations. Several socioeconomic factors related to income and education level as well as mood- and fatigue-related factors, ability to cope with stress, and energy level were consistently associated across diseases. Several occupational exposures such as dust, biohazardous material, and diesel and gasoline exhaust were also associated across diseases. Hobby exposures such as glues or paints were associated with many diseases. The following sections discuss a subset of the associations found in the ExWAS. Results were chosen primarily from questions on occupational exposures from the Health and Exposure Survey and External Exposome Survey due to the larger sample sizes of the analysis and novelty of the results. Comprehensive results can be explored in PEGS Explorer or the supplemental file provided.

Allergic rhinitis

Allergic rhinitis was associated with skin-irritating exposures such as dust [adjusted *P*-value ≤0.001, OR(95%CI) = 1.95(1.68,2.26)], chemicals/acids/solvents [<0.001, 1.36(1.22,1.52)], diesel exhaust [0.003,1.56(1.3,1.88)], gasoline exhaust [<0.001,1.51(1.26,1.80)], plastics [0.025, 1.69(1.18,2.42)], and any paints/paint thinners [0.034,1.39 (1.10,1.77)]. The results replicated known associations for many exposures. Because of the large number of exposures associated with the disease, it is difficult to evaluate causes or pathways of allergic rhinitis individually.

Asthma

Asthma was associated with occupational exposures that include dust [<0.001, 2.02(1.66,2.44)], diesel exhaust [0.003,1.53(1.17,1.97)], gasoline exhaust [0.011,1.44(1.11,1.85)], dyes [0.011,1.58(1.14,2.17)], textile fibers/dust [<0.001,1.63(1.25,2.09)], asbestos [<0.001,1.65 (1.28,2.11)], inorganic dust such as sand [0.001,1.81(1.29,2.49)], coal or stone dust [<0.001,2.19(1.47,3.18)], heavy metals [0.004,2.31 (1.51,3.45)], chemicals/acids/solvents [<0.001,1.40(1.19,1.63)], mold [0.025,1.55(1.16,2.06)], and cleaning chemicals [0.004,1.60(1.26,2.03)]. The results confirm these known associations and demonstrate the

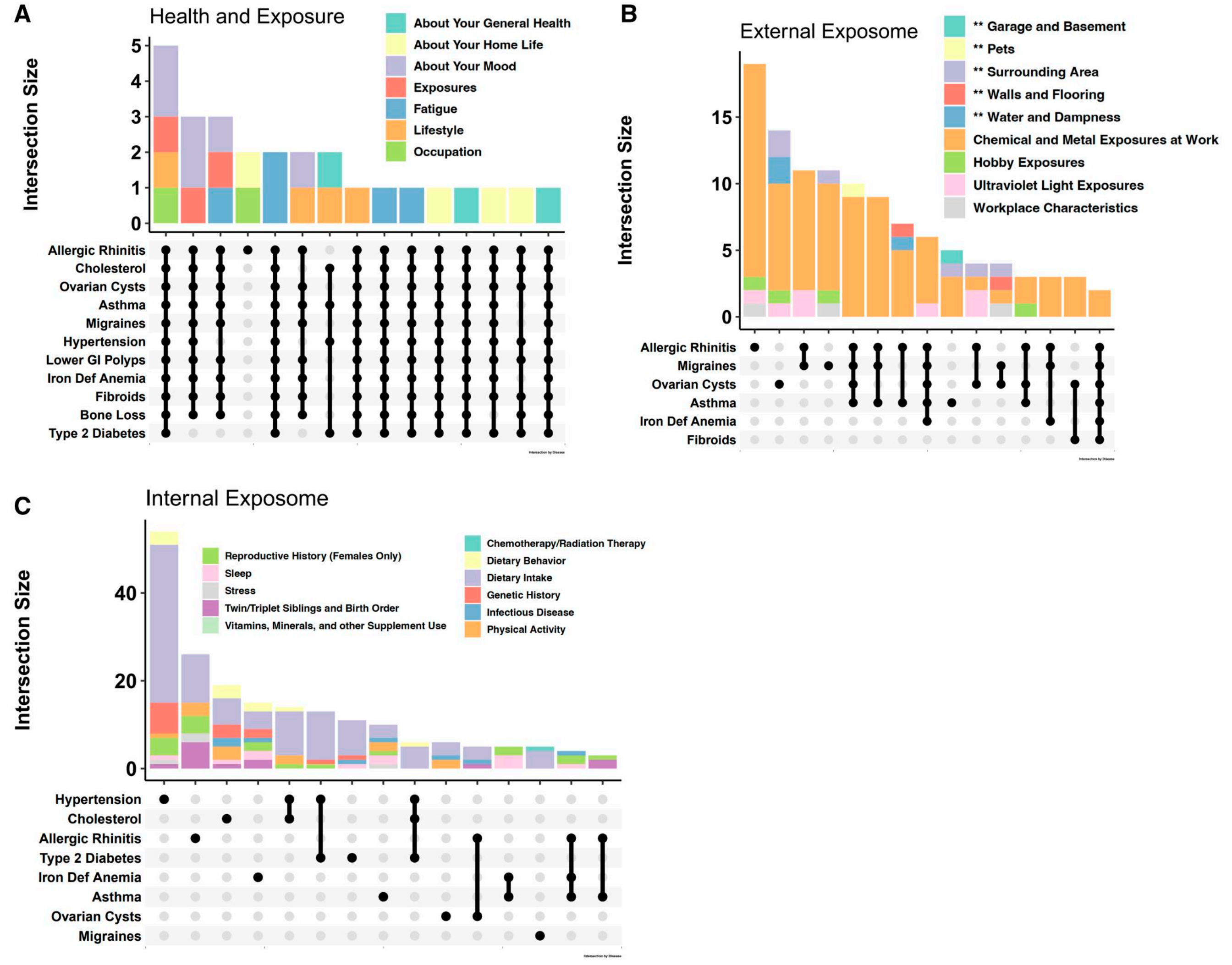

Figure 3. UpSet plots showing the intersection of common exposures across phenotypes for the (A) Health and Exposure Survey, (B) External Exposome Survey, and (C) Internal Exposome Survey. The black circles indicate which exposures were associated with other phenotypes. The bar charts show common exposures, with topical survey sections represented by the colors in the individual legends. **Characteristics of current and past residences.

variety of sources of irritants for asthma. The results also revealed new associations for asthma and textile fibers/dust [<0.001,1.63 (1.25,2.09)], acids [<0.001,1.40(1.19,1.63)], and living in a home with plastic or vinyl floors [0.054,1.37(1.09,1.73)].

Bone loss

Bone loss was associated with various dust exposures including general dust [<0.001,1.75(1.37,2.22)], coal or stone dust [0.008,2.07 (1.29,3.26)], textile fibers/dust [0.037,1.42(1.07,1.87)], and wood dust [0.012,1.66(1.18,2.32)]. Bone loss was also associated with exposure to gasoline exhaust [0.084,1.40(1.02,1.89)], pesticides/herbicides [0.082,1.39(1.02,1.88)], and dyes [0.094,1.44(1.01,2.04)]. Linkages with other occupational exposures are less clear as research on bone loss is sparser than for the other considered traits. The association of bone loss with both organic and inorganic dust replicate findings in prior work.

Fibroids

Fibroids were associated with exposure to biohazardous material [<0.001,1.38(1.18,1.61)], pesticides/herbicides [0.039,1.53(1.07,2.17)], and heavy metals [0.033,1.69(1.10,2.56)], all of which have been previously linked to fibroids. There are also associations with exposure to dust [<0.001,1.77(1.37,2.29)], textile fibers/dust [0.003,1.61 (1.21,2.14)], asbestos [0.016,1.51(1.12,2.02)], formaldehyde [0.008, 1.49(1.15,1.94)], diesel exhaust [0.005,1.86(1.26,2.73)], and gasoline exhaust [0.018,1.60(1.13,2.25)]. We also found associations with exposure to dyes [0.061,1.50(1.03,2.15)], replicating previous work associating fibroids with dyes and other hair products.

High cholesterol

High cholesterol was associated with exposure to general dust [0.001,1.49(1.27,1.74)], coal/stone dust [0.009,1.58(1.16,2.17)], asbestos [<0.001,1.55(1.28,1.87)], and wood dust [0.01,1.34(1.09,1.65)]. There is limited literature on associations between high cholesterol and dust exposure levels. High cholesterol was also associated with biohazardous materials [0.007,1.19(1.06,1.33)]. An association was also found with textile fibers/dust [<0.001,1.93(1.58,2.35] and dyes [0.001,1.57(1.22,2.03)]. Research on associations of high cholesterol with textile-related exposures is sparse, so this is a potential area for further study.

Hypertension

Hypertension was associated with exposure to biohazardous material [<0.001,1.29(1.14,1.45)] and dust [0.002,1.33(1.13,1.57)]. This confirms previous findings of associations of hypertension with particulate matter from many sources. Asbestos pleurisy has also been associated with hypertension, which is in accordance with our finding of an association with asbestos exposure [0.003,1.38(1.14,1.68)]. Accordingly, asbestos exposure may be involved in the mechanism that induces hypertension. The results replicate a previous finding of an association of hypertension with diesel exhaust [<0.001,1.54(1.26,1.89)] and gasoline exhaust [0.043,1.26(1.04,1.54)]. Hypertension was also associated with exposure to textile fibers/dust [0.037,1.29(1.05,1.59)]. There is little research on associations with textile-related exposures, so this is a potential area for further study.

Iron-deficient anemia

Iron-deficient anemia was associated with exposure to asbestos [<0.001,2.12(1.68,2.64)], chemicals/acids/solvents [<0.001,1.38 (1.19,1.60)], coal/stone dust [<0.001,2.16(1.43,3.20)], coal tar/pitch/asphalt [<0.001,3.12(1.95,4.87)], diesel exhaust [0.072,1.34(1.01, 1.76)], gasoline exhaust [0.012,1.44(1.11,1.86)], dyes [0.003,1.64 (1.21,2.19)], heavy metals [0.003,1.70(1.23,2.32)], pesticides [0.013, 1.46(1.11,1.90)], sand/silica [0.059,1.48(1.02,2.10)], general dust [<0.001,1.76(1.44,2.14)], textile fibers/dust [<0.001,1.82(1.44,2.30)], and wood dust [0.002,1.65(1.23,2.18)]. Overall, research on exposures that affect iron-deficient anemia is sparse.

Lower GI polyps

Lower GI polyps were associated with exposure to asbestos [0.049,1.33(1.04,1.69)], pesticides/herbicides [0.03,1.40(1.08,1.81)], and chemicals/acids/solvents [0.013,1.26(1.08,1.48)], confirming findings in previous work. The ExWAS results included novel associations between lower GI polyps and exposure to formaldehyde [0.059,1.31(1.03,1.66)], textile fibers/dust [0.0091.49 (1.15,1.93)], and other types of dust [<0.001,1.62(1.31,2.00)]. These results demonstrate the variety of exposures that affect lower GI polyps and indicate the need for further study.

Migraines

The association of migraines with exposure to lead through fishing with lead weights [0.04,1.60(1.14,2.21)] was unique to migraines in our ExWAS. Airborne exposures such as dust [<0.001,1.55(1.28,1.88)], textile fibers/dust [0.002,1.50(1.18,1.89)], and dyes [0.024,1.46(1.08,1.95)] were associated with migraines. Other associations were exposure to coal/stone dust [<0.001,2.00 (1.35,2.90)], coal tar [<0.001, 2.67(1.70,4.07)], formaldehyde [<0.001,1.58(1.28,1.94)], and pesticides/herbicides [<0.001,1.58 (1.22,2.03)], which may be involved in the mechanism for the development of chronic migraines.

Ovarian cysts

Ovarian cysts were associated with exposure to heavy metals [0.044,1.61(1.04,2.46)], confirming the results of prior work. Ovarian cysts were also associated with dust [<0.001,1.67 (1.29,2.16)], gasoline exhaust [0.015,1.59(1.12,2.23)], diesel exhaust [<0.001,2.12(1.47,3.03)], dyes [0.073,1.46(1.00,2.11)], formaldehyde [0.006,1.48(1.14,1.90)], and textile fibers/dust [0.006,1.59 (1.17,2.14)], none of which have been thoroughly studied. Additional findings of an association of ovarian cysts with biohazardous materials [<0.001,1.55(1.33,1.80)] and chemicals/acids/solvents [<0.001,1.57(1.33,1.86)] suggest a variety of exposures affect the development of this disease.

Type 2 diabetes

Type 2 diabetes was associated with exposure to diesel exhaust [0.007,1.52(1.15,1.99)], gasoline exhaust [0.007.1.50(1.14,1.94)], coal/stone dust [0.004,1.93(1.28,2.86)], and biohazardous material [0.064,1.22(1.01,1.47)]. Associations were also found for exposure to asbestos [0.004,1.53(1.18,1.96)] and textile fibers/dust [0.007,1.50(1.15,1.95)], confirming previously published work.

Summary

There were associations across diseases for chronic fatigue syndrome diagnosis, general fatigue, changes in mood, and sleep issues. These factors are most likely caused by the diseases of interest, as all of these factors entail an individual burden that could cause issues with sleep, fatigue, and a generally negative mood. Diagnosis with chronic fatigue syndrome may indicate a more valid association as this is a doctor-diagnosed syndrome that has major effects on daily life.

There were also associations across diseases with socioeconomic factors. The results replicate findings in prior work and demonstrate that socioeconomic factors can impact disease risk, particularly through the negative health effects of low socioeconomic status.

Diagnosis with infectious diseases such as the flu, food poisoning, sexually transmitted infections, and hepatitis were associated with a variety of phenotypes. Viral pneumonia was associated with 10 of our 11 diseases, with the only null association being with bone loss.

Family history of disease also affects health, demonstrated by the association of family history of hyperlipidemia and hypercholesterolemia for the expected phenotypes such as high cholesterol and type 2 diabetes but also for iron-deficient anemia. Further analysis of the results can be found in the PEGS Explorer website and the associated manuscript.

Causaly results

Causaly analysis enabled the discovery of novel and replicated exposures in the literature that would otherwise be difficult to locate. Figure 4 provides an overview of common PEGS exposures and indicates whether they were replicated in Causaly findings and, if so, the strength of the evidence. Supplementary Table S1 lists the search terms used. The results provide a sense of the weight of evidence for associations between exposure types and individual diseases and systematically identify existing and novel findings.

Allergic rhinitis

Results for allergic rhinitis in Causaly included occupational exposures, namely dust (number of manuscripts found = 122), diesel exhaust (23), textiles (2), and ethanol (1), and lifestyle factors, namely sleep (103), education (43), fatigue (42), and smoking (13).

Asthma

Results for asthma included lifestyle and socioeconomic factors, namely smoking (1,445), educational level (759), sleep apnea (335), general fatigue (26), and income (24). The results also included occupational exposures, namely general dust (563), asbestos (9), wood dust (1), and textile fibers (2). Phthalate, a plastic derivative (405), and diesel exhaust (135) were also included in the results. Mold and dampness (380) were associated with the development of asthma.

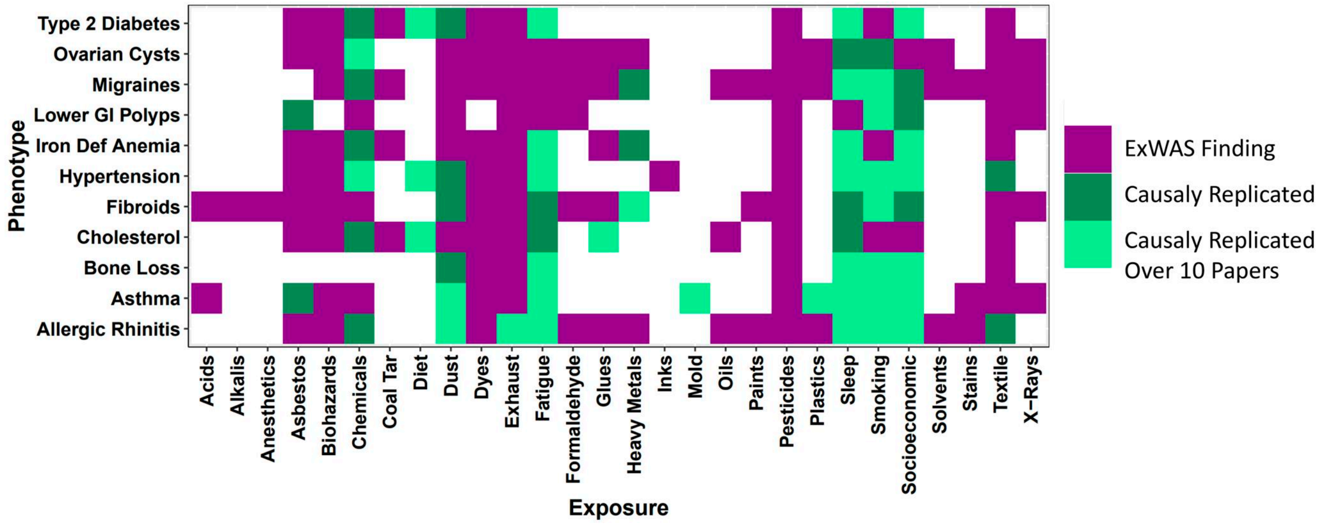

Figure 4. Tile plot for ExWAS findings and their replication in a Causaly literature search. On the x-axis are broad categories used for the Causaly literature search, and on the y-axis are the disease phenotypes. A blank square indicates that the ExWAS had null results, a purple square is a finding with no direct replication, a dark green square was replicated in a few studies, and a light green square was a highly replicated finding.

Bone loss

The results for bone loss included lifestyle factors, namely smoking (217), fatigue (84), low-quality sleep (29), and educational status (24). The results also included the occupational exposure of dust (6).

Fibroids

The results for fibroids included the heavy metals cadmium (67) and mercury (1) whereas the ExWAS included only general heavy metals exposure. The results also included smoking (16), dust (1), and glues (1). Lifestyle factors such as fatigue (1), sleep (1), and lack of employment (6) were also returned in the results.

High cholesterol

The results for high cholesterol include the lifestyle factors of diet (219), smoking (17), sleep (2), and fatigue (1). The results also included ethanol (5).

Hypertension

The results for hypertension included lifestyle factors, namely sleep apnea and other sleep disorders (2,843), smoking (1,113) educational status (1,077), diet (128), and fatigue (51). The results also included ethanol (11), chlorine, (3), biological toxins (3), textiles (6), and dust (7).

Iron-deficient anemia

The results for iron-deficient anemia included fatigue (70), poverty (26), lack of occupational activity (23), smoking (5), and lack of sleep (12). The results also included the heavy metals cadmium (3) and lead (3) as well as environmental pollutants (2), ethanol (1), and biological toxins (1).

Lower GI polyps

The results for GI polyps included the lifestyle factors of smoking (23) and general fatigue (5) and the occupational exposure of asbestos (1).

Migraines

The results for migraines included the lifestyle factors of lack of sleep (381), smoking (55), and sleep disorders (43). The results also included educational status (9) and occupational exposures, namely ethanol (6) and lead (1).

Ovarian cysts

The results for ovarian cysts were ethanol (18), smoking (1), and sleep (1).

Type 2 diabetes

The results for type 2 diabetes included lifestyle exposures, namely diet (6,947), sleep disorders (921), sleep issues (38), and fatigue (31), and the socioeconomic factor of income (21). The results also included occupational exposures, namely ethanol (3), dust (2), environmental pollutants (1), asbestos (1), and coal dust (1).

Summary

Causaly helped establish which associations found in the ExWAS results are novel and which were previously established. We determined the relative strength of the association by considering the number of manuscripts associated with each exposure. The results of the present study will add to the strength of the associations. We found established relationships from our previous studies in the Casualy database, giving us confidence that we are capturing current research on the associations between exposures and disease status.

Deletion/substitution/addition (DSA) results

The DSA results provide a comprehensive view of how exposures interact for each trait. A single-exposure model ignores correlations between exposures and assumes that exposures are linear and do not occur concurrently. In contrast, a DSA model allows the simultaneous, iterative consideration of multiple exposures to produce a single model. Exposures recorded by all surveys were used as input for the DSA models for the considered diseases. Supplementary Table S2 summarizes the results of the DSA analysis.

Discussion

Looking at the totality of exposures across multiple phenotypes helps engender a holistic understanding of how the exposome affects health and disease status and elucidates which exposures are associated with multiple traits. The ExWAS results reveal associations between exposures from multiple facets of an individual's life and how they are associated with multiple diseases. The results for additional sensitivity analyses stratified by sex are available in PEGS Explorer. The results from stratified analyses can be used to generate important biological hypotheses, as they reveal how women and men experience and are affected differently by exposures.

An important aspect of this study is the tools we concurrently developed to interpret these results, with the goal of expanding understanding of the exposome and its effects on human health. These tools also demonstrate the data that are available and how they can be analyzed. These tools are discussed in a related paper in this issue focused on the PEGS Explorer web application.

The ExWAS results revealed both novel and established associations of many exposures with multiple diseases, suggesting that common environmental factors influence disease status across phenotypes. An interesting finding is the association of textile-related exposures with most of the considered diseases. While it is not possible to establish causality from the results, this suggests an underlying association between occupations involving textile fiber exposure and these diseases. Another interesting finding is the consistent association of asbestos with the considered phenotypes. Further, the consistent association of dust exposure and multiple phenotypes indicates that further investigation of this association is warranted. While this association is well-studied for some diseases, including asthma, there is scant work investigating its effects on diseases that include fibroids and lower GI polyps and how dust components affect multiple biological systems and pulmonary function. Overall, disseminating the ExWAS results will help understand the mechanisms by which common exposures affect disease status.

Integrating additional data types such as geospatial data and environmental measurements with questionnaire-based exposome data will help researchers discover co-occurring exposures and could offer insights into how some populations are disproportionately affected by some exposures. There are known differences in both trait prevalence and levels of exposure across strata, including by race/ethnicity and sex,[21,32,33] as our sensitivity analyses in the PEGS cohort demonstrate. While it is not possible to infer causal relationships from ExWAS in survey data, it is possible to determine whether the exposures associated with various traits and how the strength of these associations changes by these strata. Follow-up with functional analysis may provide insights into which occupational exposures are the most strongly associated with specific traits, with the potential to determine causality.

This study has several strengths. First, the exposure data collected by PEGS is broader compared to other epidemiological studies and includes diet, exposure to chemical/biological agents at home and work, childhood socioeconomic status, family medical history, and psychosocial stressors. These questionnaire-based internal and external exposure data are a valuable resource for better understanding how exposure and lifestyle factors are associated with each other and with multiple traits and disease in general. Second, ExWAS enables researchers to determine associated exposures for functional follow-up. Conducting functional follow-up is expensive and can result in fatigue for participants called for follow-up, so it is important to use effective methods for determining candidate exposures. ExWAS results can help researchers narrow down which exposures should be the focus of follow-up. For example, because of the significant association of diesel exhaust with fibroids, follow-up could examine the mechanism or pathways through which diesel exhaust affects the development of fibroids.

Third, the results provide insights into exposures associated with multiple traits, also helping to identify candidates for further study. For example, exposure to diesel and gasoline exhaust and asbestos are known to be harmful, and their association with multiple traits suggests these exposures are candidates for replication in follow-up work. Fourth, this study is one of only a few to examine exposures across diseases to identify environmental risk factors that may explain the pathophysiology and complex etiology of multiple traits. The use of multi-exposure models underscores the importance of considering concurrent exposures.

To avoid misinterpretation, it is important to consider the limitations of this work. First, some case groups have small sample sizes, resulting in imprecise estimates with wide confidence intervals for groups with $n < 10$. Second, as PEGS is an observational study and exposure data were collected cross-sectionally, differential recall may be an issue. Additionally, the PEGS cohort comprises over 60% women, and certain exposures may be more or less prevalent in the cohort than in the larger population. Further, the sampling methods used in PEGS may lead to complications with temporality as exposure surveys were sometimes conducted after disease diagnoses, potentially leading to length-biased sampling. This can create issues with reverse causality when identifying predictive risk factors. Third, there is a possibility that exposome data were misclassified. However, as all participants completed the same survey regardless of their health status, the authors assume non-differential misclassification between subjects, leading to potential bias toward the null.

While self-reported data may be affected by information bias, participants were asked whether their health status/trait was diagnosed by a healthcare provider, potentially minimizing this bias. Next, the ExWAS results were used to generate hypotheses about underlying mechanisms and study associations, and it is important to replicate these associations in an independent sample.

Finally, the limited sample size of the Internal and External Surveys, which were completed by only a subset of initial study participants, may have caused important associations to be missed due to lack of power. This lack of power was also due to the exclusion of questions dependent on responses to other questions because of the small sample sizes for individual response levels. Accordingly, although our analyses included broad exposure information, it was difficult to focus on specific elements of some exposures.

Providing methods to analyze exposome data from the diverse PEGS cohort and disseminating the ExWAS results supports a collaborative approach to in-depth analysis of the exposome and the examination of the association of exposures with multiple phenotypes. Moving exposomics toward a more collaborative space such as that in genomics will increase knowledge of the field and support the replication and accuracy of results across cohorts.

Acknowledgments

We wish to thank Hannah Collins Cakar for assistance with manuscript submission.

Supplementary data

Supplementary material is available at *Exposome* online.

Funding

Financial support was received from intramural funds from the National Institutes of Health, National Institute of Environmental Health Sciences.

Author contributions

Dillon Lloyd (Data curation [equal], Formal analysis [equal], Writing—original draft [equal]), John S. House (Data curation [equal], Project administration [equal], Writing—review & editing [equal]), Farida S. Akhtari (Data curation [equal], Writing—review & editing [equal]), Charles Schmitt (Data curation [equal], Writing—review & editing [equal]), David C. Fargo (Data curation [equal], Writing—review & editing [equal]), Elizabeth H. Scholl (Data curation [equal], Formal analysis [equal], Visualization [equal], Writing—review & editing [equal]), Jason Phillips (Data curation [equal], Software [equal], Visualization [equal], Writing—review & editing [equal]), Shail Choksi (Software [equal], Visualization [equal], Writing—review & editing [equal]), Ruchir Shah (Software [equal], Visualization [equal], Writing—review & editing [equal]), Janet E. Hall (Funding acquisition [equal], Writing—review & editing [equal]), and Alison A. Motsinger-Reif (Conceptualization [equal], Funding acquisition [equal], Project administration [equal], Writing—review & editing [equal]).

Conflict of interest statement

None declared.

References

1. Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol Biomarkers Prev. 2005;14(8):1847–1850. doi:10.1158/5–9965.Epi-5–0456
2. Wild CP. The exposome: from concept to utility. Int J Epidemiol. 2012;41(1):24–32. doi:10.1093/ije/dyr236
3. Kolb H, Martin S. Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes. BMC Med. 2017;15(1): 131.doi:10.1186/s12916-7–0901-x
4. Miller GW, Jones DP. The nature of nurture: refining the definition of the exposome. Toxicol Sci. 2013;137(1):1–2. doi: 10.1093/toxsci/kft251
5. Yarnell J, Yu S, McCrum E, *et al.* Education, socioeconomic and lifestyle factors, and risk of coronary heart disease: the PRIME Study. Int J Epidemiol. 2004;34(2):268–275. doi:10.1093/ije/dyh267
6. Kampa M, Castanas E. Human health effects of air pollution. Environ Pollut. 2008;151(2):362–367. doi:10.1016/j.envpol.2007.06.012
7. Thiering E, Heinrich J. Epidemiology of air pollution and diabetes. Trends Endocrinol Metab. 2015;26(7):384–394. doi:10.1016/j.tem.2015.05.002
8. Gold DR, Samet JM. Air pollution, climate, and heart disease. Circulation 2013;128(21):e411–e414. doi:10.1161/CIRCULATIONAHA.113.003988
9. Halder JN, Islam MN, Department of Environmental Engineering, Sangji University, Republic of Korea. Water pollution and its impact on the human health. J Environ Hum 2015;2(1):36–46. doi: 10.15764/EH.2015.01005
10. Akhtari FS, Lloyd D, Burkholder A, *et al.* Questionnaire-based polyexposure assessment outperforms polygenic scores for classification of type 2 diabetes in a multiancestry cohort. Diabetes Care. 2023;46(5):929–937. doi:10.2337/dc22-0295
11. Lee EY, Akhtari F, House JS, *et al.* Questionnaire-based exposome-wide association studies (ExWAS) reveal expected and novel risk factors associated with cardiovascular outcomes in the Personalized Environment and Genes Study. Environ Res. 2022;212(Pt D):113463.doi:10.1016/j.envres.2022.113463
12. Hu H, Zhao J, Savitz DA, *et al.* An external exposome-wide association study of hypertensive disorders of pregnancy. Environ Int. 2020;141:105797.doi:10.1016/j.envint.2020.105797
13. Cui Y, Balshaw DM, Kwok RK, *et al.* The exposome: embracing the complexity for discovery in environmental health. Environ Health Perspect. 2016;124(8):A137–A140. doi:10.1289/EHP412
14. Agache I, Miller R, Gern JE, *et al.* Emerging concepts and challenges in implementing the exposome paradigm in allergic diseases and asthma: a Practall document. Allergy 2019;74(3): 449–463. doi:10.1111/all.13690
15. Chung MK, Louis GMB, Kannan K, Patel CJ. Exposome-wide association study of semen quality: Systematic discovery of endocrine disrupting chemical biomarkers in fertility require large sample sizes. Environ Int. 2019;125:505-514. doi:10.1016/j.envint.2018.11.037
16. Haddad N, Andrianou X, Parrish C, Oikonomou S, Makris KC. An exposome-wide association study on body mass index in adolescents using the National Health and Nutrition Examination Survey (NHANES) 2003–2004 and 2013–2014 data. Sci Rep. 2022; 12(1):8856.doi:10.1038/s41598-2–12459-z
17. Vrijheid M, Fossati S, Maitre L, *et al.* Early-life environmental exposures and childhood obesity: an exposome-wide approach. Environ Health Perspect. 2020;128(6):67009.doi:10.1289/ehp5975
18. Noonan AS, Velasco-Mondragon HE, Wagner FA. Improving the health of African Americans in the USA: an overdue opportunity for social justice. Public Health Rev. 2016;37:12.doi:10.1186/s40985-6–0025-4
19. Cifuentes P, Reichard J, Im W, *et al.* Application of the public health exposome framework to estimate phenotypes of resilience in a model Ohio African-American women's cohort. J Urban Health. 2019;96(Suppl 1):57–71. doi:10.1007/s11524-8–00338-w
20. Juarez PD, Hood DB, Song M-A, Ramesh A. Advancing the Science of Cancer in Latinos: Building Collaboration for Action. Springer International Publishing Cham; 2022:7–32.
21. Juarez PD, Hood DB, Song MA, Ramesh A. Use of an exposome approach to understand the effects of exposures from the natural, built, and social environments on cardio-vascular disease onset, progression, and outcomes. Front Public Health. 2020;8: 379.doi:10.3389/fpubh.2020.00379
22. Ayala-Ramirez M, MacNell N, McNamee LE, *et al.* Association of distance to swine concentrated animal feeding operations with immune-mediated diseases: An exploratory gene-environment study. Environ Int. 2023;171:107687.doi:10.1016/j.envint.2022.107687
23. Lowe ME, Akhtari FS, Potter TA, *et al.* The skin is no barrier to mixtures: air pollutant mixtures and reported psoriasis or eczema in the Personalized Environment and Genes Study (PEGS). J Expo Sci Environ Epidemiol. 2023;33(3):474–481. doi:10.1038/s41370-2–00502-0
24. Lee EY, Choi W, Burkholder AB, *et al.* Race/ethnicity-stratified fine-mapping of the MHC locus reveals genetic variants associated with late-onset asthma. Front Genet. 2023;14:1173676.doi:10.3389/fgene.2023.1173676

25. Patel CJ, Bhattacharya J, Butte AJ. An environment-wide association study (EWAS) on type 2 diabetes mellitus. PLoS One. 2010; 5(5):e10746.doi:10.1371/journal.pone.0010746
26. Patel CJ. Introduction to environment and exposome-wide association studies: a data-driven method to identify multiple environmental factors associated with phenotypes in human populations. In: Rider CV, Simmons JE, eds. Chemical Mixtures and Combined Chemical and Nonchemical Stressors: Exposure, Toxicity, Analysis, and Risk. Springer International Publishing; 2018:9–149.
27. Braun JM, Kalloo G, Kingsley SL, Li N. Using phenome-wide association studies to examine the effect of environmental exposures on human health. Environ Int. 2019;130:104877.doi:10.1016/j.envint.2019.05.071
28. Avery CL, Howard AG, Ballou AF, *et al.* Strengthening causal inference in exposomics research: application of genetic data and methods. Environ Health Perspect. 2022;130(5):55001.doi:10.1289/EHP9098
29. Manrai AK, Cui Y, Bushel PR, *et al.* Informatics and data analytics to support exposome-based discovery for public health. Annu Rev Public Health. 2017;38:279–294. doi:10.1146/annurev-publhealth-082516-012737
30. Bernhard MC, Evans MB, Kent ST, *et al.* Identifying environmental health priorities in underserved populations: a study of rural versus urban communities. Public Health. 2013;127(11): 994–1004. doi:10.1016/j.puhe.2013.08.005
31. Morrone M, Tres A, Aronin R. Creating effective messages about environmental health. J Environ Health. 2005;68(1):9.
32. Juarez PD, Matthews-Juarez P, Hood DB, *et al.* The public health exposome: a population-based, exposure science approach to health disparities research. Int J Environ Res Public Health. 2014;11(12):12866–12895. doi:10.3390/ijerph111212866
33. Sum KK, Tint MT, Aguilera R, *et al.* The socioeconomic landscape of the exposome during pregnancy. Environ Int. 2022;163: 107205. doi:10.1016/j.envint.2022.107205