
AFRL-RH-WP-TR-2020-0102

# Review of Open Access and DoD Biomarker-Related Databases for Exposure Assessment Research

**Brett Edwards and Jennifer Schwanekamp Kerr**
**UES, Inc.**

**Logan Hyder**
**Henry M Jackson Foundation**

**Christin Duran**
**711 HPW/RHBC**

**NOVEMBER 2020**
**Final Report**



**AIR FORCE RESEARCH LABORATORY**
**711TH HUMAN PERFORMANCE WING,**
**AIRMAN SYSTEMS DIRECTORATE,**
**WRIGHT-PATTERSON AIR FORCE BASE, OH 45433**
**AIR FORCE MATERIEL COMMAND**
**UNITED STATES AIR FORCE**

# NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th Air Base Wing Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RH-WP-TR-2020-0102 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

| | |
|---|---|
| //signature// | //signature// |
| CHRISTIN M. DURAN, PhD, CIH | LOGAN A. WILLIAMS, PhD |
| Site lead, Research Chemical Engineer | Core Research Area Lead |
| Performance Optimization Branch | Performance Optimization Branch |
| Airman Biosciences Division | Airman Biosciences Division |

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE *(DD-MM-YY)* | 2. REPORT TYPE | 3. DATES COVERED *(From - To)* |
|---|---|---|
| 11-25-2020 | Final | 08/2019 – 08/2020 |

**4. TITLE AND SUBTITLE**
Review of Open Access and DoD Biomarker-Related Databases for Exposure Assessment Research

**5a. CONTRACT NUMBER**
FA8650-17-C-6834

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
*Brett Edwards, Jennifer Schwanekamp-Kerr
**Logan Hyder
***Christin Duran

**5d. PROJECT NUMBER**
18-190.1

**5e. TASK NUMBER**
00

**5f. WORK UNIT NUMBER**
Legacy RHM

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
*UES, Inc.
4401 Dayton Xenia Road
Dayton, OH 45432
**Henry M. Jackson Foundation

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
***Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Airman Systems Directorate
Airman Biosciences Division
Wright-Patterson AFB, OH 45433

**10. SPONSORING/MONITORING AGENCY ACRONYM(S)**
711 HPW/RHBC

**11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)**
AFRL-RH-WP-TR-2020-0102

**12. DISTRIBUTION/AVAILABILITY STATEMENT**
Distribution A. Approved for public release.

**13. SUPPLEMENTARY NOTES**
AFRL-2020-0461, cleared on 8 December 2020

**14. ABSTRACT**

The 711th HPW requires the ability to rapidly respond to military exposure events with recommendations for measurable biomarkers of exposure and effect. Over the past 10 years, publications related to biomarkers of exposure have increased, which coincides with a paradigm shift to a more individualistic exposure assessment where biological burden and genetic factors contributing to susceptibility are included. This report is a review of a subset of open access databases and Department of Defense databases. In total, ten open access databases were evaluated and tested for utility using three military exposures of interest: copper, particulate matter, and organophosphates. The Comparative Toxicogenomics Database provided the most comprehensive and useful results by documenting biological specimens collected, concentrations of chemicals and markers in samples, and citations for corresponding literature. Six Department of Defense databases were selected for analysis in this report. Our analysis found that these databases provide information related to individual warfighter health and exposure but do not adequately identify measurable biomarkers and methods for detection that can be applied in exposure assessments. The move towards comprehensive individual exposure assessment requires synthesis of data from multiple databases to inform new and updated exposure assessment protocols for military relevant exposures.

**15. SUBJECT TERMS**
Biomarker, biomonitoring, database, exposure

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT: | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON (Monitor) |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Dr. Christin Duran |
| Unclassified | Unclassified | Unclassified | SAR | 38 | 19b. TELEPHONE NUMBER *(Include Area Code)* N/A |

**TABLE OF CONTENTS**

## LIST OF FIGURES



## LIST OF TABLES

## 1.0 INTRODUCTION

Exposure monitoring and assessment is important to ensure the health and safety of workers. In 1991 the Department of Defense (DoD), in response to health impacts of the Gulf War, initiated a risk assessment model to monitor occupational and environmental health of military during deployments (1). Traditionally, DoD exposure assessment of occupational and environmental exposures has relied on air, soil, and water measurements, and self-reporting via questionnaires (1, 2). More recently, exposure assessment monitoring approaches have expanded to include the use of individual biomonitoring to identify important exposures (2, 3). The goal of these changes is to develop a hazard assessment system to identify the source, internal dose, and health consequences of exposure.

Broadly, biomarkers are defined as any reliable characteristic measured in the body that is indicative of normal biological processes, pathogenic processes, or response to exposure and/or therapeutic intervention (3-5). Biomarkers can be further categorized as biomarkers of exposure, effect, or susceptibility (3-6). Biomarkers of exposure are chemicals and/or metabolites in biofluids, such as blood, urine or saliva that are specific to the exposure (3, 6, 7). Biomarkers of effect are biochemical, physiological, or behavioral changes associated with an exposure or a resulting disease state (3, 6, 7). Finally, biomarkers of susceptibility are typically genetic factors or pre-existing diseases that may impair how an individual responds to an exposure (3, 6, 7) (Figure 1). Importantly, biomarkers of effect and susceptibility, including single nucleotide polymorphisms or gene expression modifications, identify subclinical changes and individuals at risk, respectively but cannot be used as direct measures of exposure (3, 6, 7).

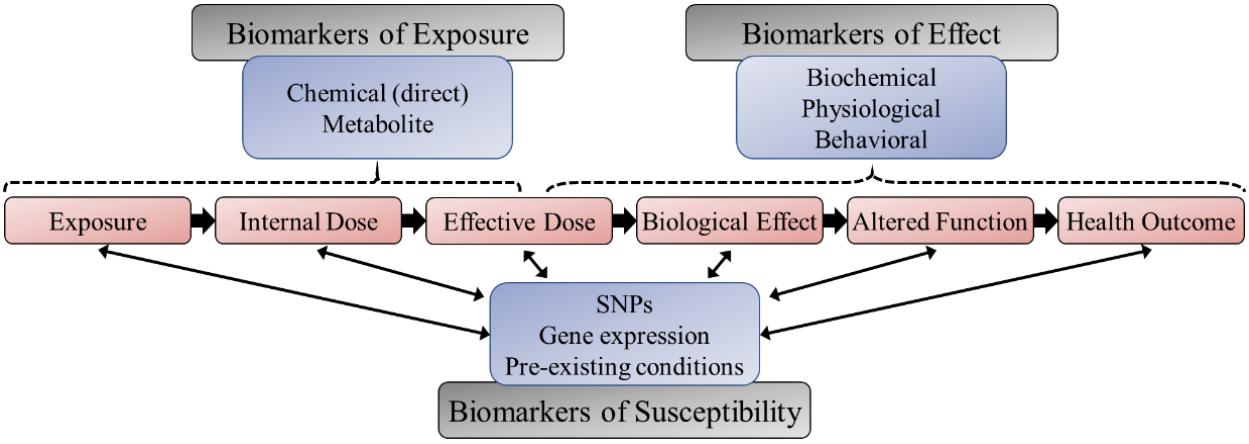



**Figure 1. Relationship between exposure and disease with respect to biomarker categories**

SNP = single nucleotide polymorphism. Adapted from (6, 8, 9).

In order to adequately and rapidly respond to exposure incidents, it is critical that response teams know which biological fluids to collect and which biomarkers to measure. The scope of existing literature on biomarkers relevant to military exposures is important to understand. In addition to literature, there are number of exposure databases publicly available where information about chemicals, exposures, relevant biomarkers, populations, and disease outcomes are synthesized from literature and other open sources. There also closed databases

specific to the DoD that are commonly used to document epidemiological and health data from military members. The goal of this report is to identify and review literature, open access databases, and DoD databases containing chemical and exposure information and evaluate their utility for informing biological measurements during exposure assessments.

## 2.0 APPROACH

### 2.1 Literature Search

To understand the scope of biomarker research in the military, a search was carried out for publications using the PubMed database (https://pubmed.ncbi.nlm.nih.gov/). Preliminary article searches were categorized using the following search criteria: "biomarkers AND military". These results were further divided into seven subgroups by utilizing tertiary search criteria ("biomarkers AND military AND cancer", "biomarkers AND military AND heart", "biomarkers AND military AND brain injury", "biomarkers AND military AND post-traumatic stress disorder (PTSD)", "biomarkers AND military AND mental health", "biomarkers AND military AND deployment", "biomarkers AND military AND exposure"). The results of these searches were filtered by species and only human reports were included in the analysis. Sources considered included books and documents, clinical trials, meta-analysis reports, randomized controlled trials, reviews, systematic reviews, and primary literature. Although the results were treated as individual publications for this analysis, it is possible that some of the publications may have overlapped in the subgroups (cancer, heart, brain injury, PTSD, mental health, and deployment). This search was completed on September 14th 2020.

### 2.2 Open Access Database Search

A search was conducted to identify open access databases containing biomarkers of exposure that can be measured in biological fluids to inform exposure assessment response protocols. The centers for disease control and prevention (CDC), the World Health Organization (WHO), and the military health system were some of the resources used to identify databases. Furthermore, some database sites also contained links to similar databases, which were also included in the review. Database information including website and notes are represented in Table 1.

Requirements for a database to meet the needs of the 711th Human Performance Wing (HPW) include documentation of exposure, biomarkers of exposure, chemical information, metabolites, biofluids used for detection, methods for measurements, physiological and/or health effects associated with the exposure, and gene and protein targets of the exposure. After a list of databases was compiled, the information provided from each database was noted in a comprehensive table. Finally, to test the utility of the database, three exposures of interest were queried: organophosphates, particulate matter, and copper. If the search term "particulate matter" did not produce results, PM10, and PM2.5 were also used. The results from the queries were compared to determine which, if any, database contained the complete set of information needed.

### 2.3 Department of Defense Database Search

DoD databases were identified utilizing three strategies. First, informational interviews were conducted with 711th HPW researchers and military personnel regarding their recommendations of databases potentially relevant to the research question. Second, the official website of the military health system (https://www.health.mil/) was used to find additional information related to data sources used previously by the 711th HPW and other operationally-

focused studies of health outcomes or exposures. This website has a reference center where fact sheets can be found highlighting capabilities of select military related databases. Finally, presentations at the 2018 Military Health System Research Symposium (MHSRS) led to the discovery of the Millennium Cohort Study database.

## 3.0 RESULTS

### 3.1 Literature Search

Based on PubMed search results, approximately 6,400 records were associated with biomarkers and the military. Evaluating these records by year, it was identified that there was a substantial increase in publications related to biological monitoring in the military between 1990 and 2019 (Figure 2, green bars). This was likely fueled by the paradigm shift to advancements in biomarker assessments tools and knowledge and more individualized research efforts over the past 30 years. On the other hand, the portion of these military biomarker publications that also contained the key word "exposure" were limited (Figure 2, blue bars). Interestingly, the total number of publications returned for both searches declined slightly from 2016 to 2019. The cause for this decline is not immediately transparent. In 2020, there was a very significant decline in publications. This search was completed on September 14th 2020, so the final number is expected to be higher than shown, but not likely as high as in 2019.

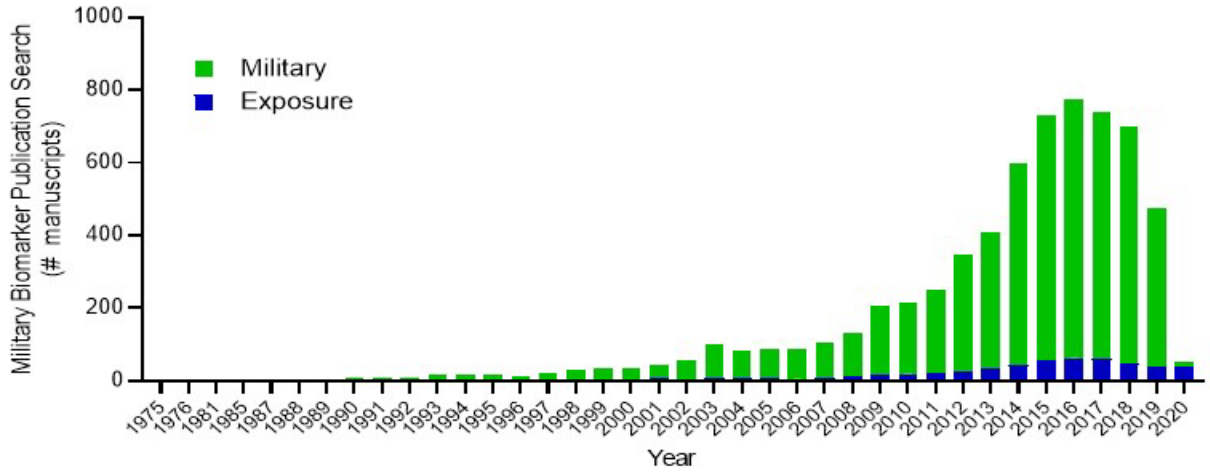



**Figure 2. Human biomarker related studies referencing the military**

PubMed search query: "biomarkers AND military" and "biomarkers AND military AND exposure". Exposure results are graphed on top of all military results to represent relative publication numbers. This search was completed on September 14th 2020.

Literature search results were filtered by several additional search criteria, including cancer, heart, brain injury, PTSD, mental health, and deployment (Figure 3). Considering each search result as an individual source, the combined subgroup (Tier 2) publications accounted for 73.6% of the military biomarker publications. The remaining 27.4% (or approximately 1,690 publications) included hepatitis, sepsis and shock, diabetes, liver and kidney disorders, and neurological disorders (Appendix, Table A1). Interestingly, over half (54%) of the military biomarker publications were related to cancer. In fact, military biomarker publications related to cancer vastly outnumber any other subgroup searched.

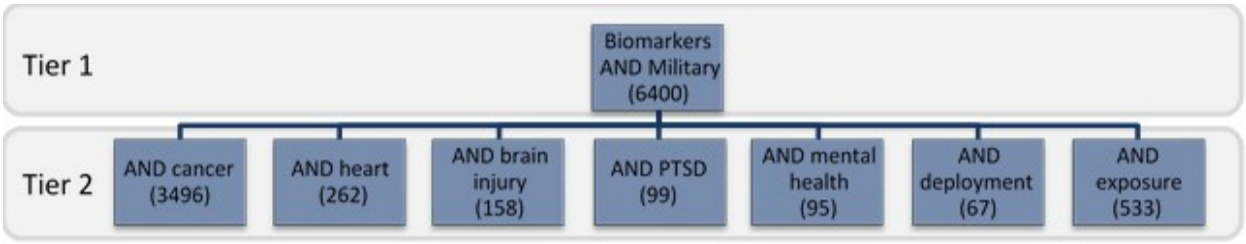



**Figure 3. Breakdown of publications by search query in PubMed**

All searches included "biomarkers AND military". The second row (Tier 2) represents subgroups differentiated by the tertiary search term. Number in parentheses = number of publications.

The results shown in Figure 3 were plotted over time to evaluate the evolution of biomarker research with respect to each subgroup (Figure 4). The biomarkers that were not associated with any of the top seven subgroup search terms were labelled "military" and are shown in grey. Publications related to biomarkers associated with cancer first appeared in 1976 and was most prevalent during most years, except for 2019. (Figure 4, blue bars). The diversity in subgroups associated with different health outcomes increased from 1998 to present. Publications associated with heart first appeared in 1998, brain injury first appeared in 1999, PTSD first appeared in 2000, deployment first appeared in 2001, and metal health first appeared in 2007.

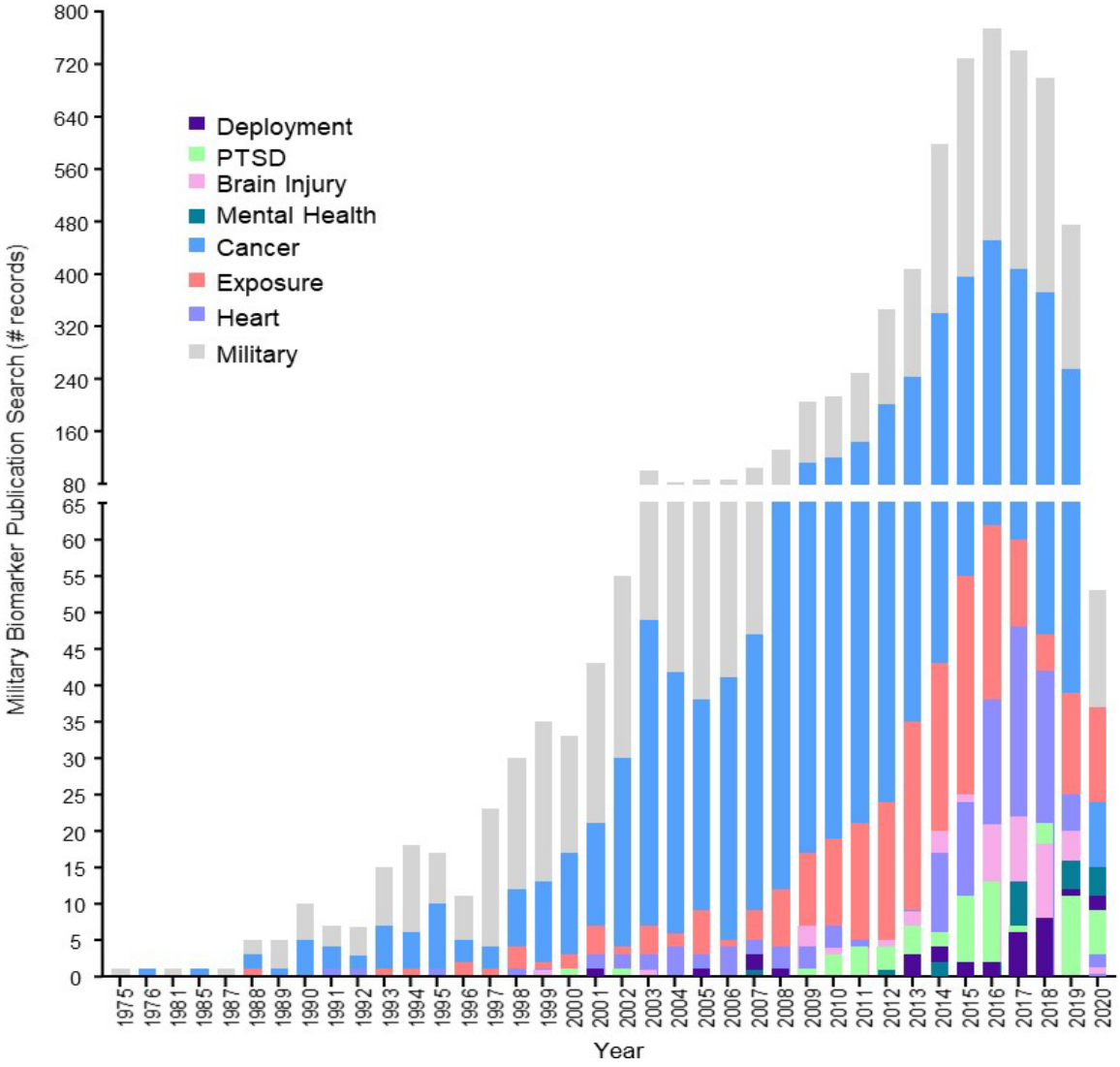



**Figure 4. Relative military related biomarker publications by year**

Military biomarker publications are represented in gray. All other searches are subcategories of military biomarker publications differentiated by the tertiary search term. Results are graphed in overlapping format to appreciate relative publication number by year. This search was completed on September 14th 2020.

## 3.2 Open Access Databases

Extensive searches revealed ten open access databases that could be useful for biomarker research. Databases detailed in this report include those specializing in disease inheritance Online Mendelian Inheritance in Man (OMIM), drug interactions (DrugBank), chemical properties (PubChem, InCHEM, ChEBI, INCHEM), metabolomics (HMBD), exposure (Exposome-Explorer, CDC National Biomonitoring Program), and Comparative Toxicogenomics Database (CTD). Information regarding access to the open access databases reviewed in this report can be found in Table 1. Additional information related to database maintenance and updates can be found in the Appendix (Table A2).

**Table 1. Information Related to Open Access Databases**

| Database | Website | Notes |
|---|---|---|
| Human Metabolome Database (HMDB) | https://hmdb.ca/ | Provides metabolites associated with search term. |
| DrugBank | https://www.drugbank.ca/ | Useful for specific metals and gases. |
| PubChem | https://pubchem.ncbi.nlm.nih.gov/ | Good for specific chemicals. |
| Chemical Entities of Biological Interest (ChEBI) | https://www.ebi.ac.uk/chebi/init.do | Good for specific chemical information. |
| Online Mendelian Inheritance in Man (OMIM) | https://omim.org/ | Genetic disease database, good for chronic diseases; can find diseases associated with exposure based on literature. |
| Exposome-explorer | http://exposome-explorer.iarc.fr/biomarkers | If biomarker is known, can search and get reference values in literature. Good for specific metabolites. |
| CDC National Biomonitoring Program | https://www.cdc.gov/biomonitoring/index.html | Limited list of biomonitoring summaries. Also contains link to chemical fact sheet. |
| International Peer Reviewed Chemical Safety Information (InCHEM) | http://www.inchem.org/#/search | Links to comprehensive reports on exposures. |
| MEB-Labs database | http://www.eh3.uc.edu/meb/ | Limited list of search terms. |
| Comparative Toxicogenomics Database (CTD) | http://ctdbase.org/ | Specific tabs for exposure studies and details. |

These databases collect source material from a variety of locations including the CDC, the National Institute of Health (NIH), PubMed, the Environmental Protection Agency (EPA), PubChem, and National Center for Biotechnology Information (NCBI) (see Appendix, Table A2).

As expected, variable results were obtained when testing the databases by searching exposures and chemicals of interest, including copper, organophosphates, and particulate matter (Table 2). Of the databases reviewed, 90% (9 out of 10) contained search pages for organophosphates and copper. However, only 35% (3.5 out of 10) contained search pages for (or

related) to particulate matter. One database had a record that indirectly mentioned particulate matter, hence a score of 0.5.

**Table 2. Results from Test Queries of Databases**

| | Query page available | | |
|---|---|---|---|
| **Database** | *Copper* | *Organophosphates* | *Particulate Matter* |
| HMDB | yes | no | indirect |
| DrugBank | yes | no | no |
| PubChem | yes | yes | yes |
| ChEBI | yes | yes | no |
| OMIM | yes | yes | no |
| Exposome-explorer | no | yes | no |
| CDC National Biomonitoring Program | no | yes | no |
| INCHEM | yes | yes | yes |
| MEB-Labs database | yes | yes | no |
| CTD | yes | yes | yes |

Indirect indicates that the database query linked to information from another database.

Additional useful exposure resources not included in detail in this report are the Occupational Safety and Health Administration (OSHA), National Institute for Occupational Safety and Health (NIOSH), and American Conferences for Governmental Industrial Hygiene (ACGIH) websites. OSHA provides enforceable permissible exposure limits that must be adhered to in order to protect employee health. NIOSH provides recommended exposure limits by considering all available medical, biological, chemical, and engineering information. Additionally, NIOSH has a searchable pocket guide to chemical hazards (https://www.cdc.gov/niosh/npg/) where chemical information, measurement methods, exposure routes, symptoms, target organs, and both OSHA and NIOSH exposure limits can be found (10). Finally, the ACGIH releases recommendations for threshold limit values (TLV) and biological exposure indices (BEI) every year. TLVs are related to what a person may be exposed to day after day without adverse health consequences. BEIs are guidance values for biological monitoring specifically relating to concentrations of chemicals that can be detected in bodily fluids. BEIs provide guidance about when to collect biological samples to make meaningful measurements (7). It is important to note that both OSHA and NIOSH exposure limit information is publicly available. However, because ACGIH is a private, non-governmental corporation, their TLVs and BEIs must be purchased.

### 3.2.1. Human Metabolome Database (HMDB)

HMDB is a metabolomics database supported by the Canadian Institute of Health Research, Canada Foundation for Innovation, and The Metabolomics Innovation Center (11). This database is designed to provide detailed information about small molecule metabolites found in the body and is intended to be used for metabolomics, clinical chemistry, biomarker discovery, and general education. This database allows the user to search by metabolites, diseases, pathways, proteins, and reactions. If the metabolite or biomarker of interest is known, the user is able to filter results by specific biospecimens, including blood, urine, saliva, and sweat. It is more challenging to search by exposure in this database.

Test queries of organophosphates, particulate matter, and copper revealed variable results, where copper was the only search term directly identified in the database (see Table 2). Organophosphates were not found in the database, nor was the more specific query, tricresyl phosphate. The phrase "particulate matter" was indirectly identified in the results for levoglucosan. Interestingly, levoglucosan is produced from the burning of carbohydrates and can be measured in urine as a biomarker for wood smoke exposure, of which particulate matter is a large component (12).

Copper does contain a specific results page that begins with a brief description. Copper is a natural ion in the body and is both an electron acceptor and donor which is important in oxidation-reduction (redox) reactions. Routes of copper exposure include dermal, inhalation, and ingestion. Important for biomonitoring applications, copper can be measured in blood, saliva, and urine. The database divides copper concentrations into normal and abnormal bins. There are 13 linked reports where copper was measured in adults and the overall concentration was between 12-22 micromolar (μM). Abnormal concentrations in blood are correlated with diseases including Wilson's disease and neurodegenerative diseases. Interestingly, for patients with Wilson's disease, a disorder in which copper builds up in the body, copper levels actually decreased in serum according to several reports.

If the metabolite of interest is known, this database is useful to obtain information on basic physiological properties, biofluids used for measurement, and concentrations detected in biofluids from published reports. However, the exposure utility of this database is limited as represented by the fact that organophosphates and particulate matter did not reveal specific results pages.

### 3.2.2. DrugBank

DrugBank is a freely available database developed and maintained by the Wisehart Research Group that combines detailed drug data with drug targets (13). Drug information includes chemical, pharmacological, pharmaceutical, and toxicological data. Drug targets include enzymes, carriers, and transporters. While this database is not specific to exposure, it could be used to identify biological changes that may occur due to confounding factors, such as medications or drug interactions. When examining occupational exposure effects, it is important to consider how medications (prescription and over the counter) may impact how the body responds. In fact, medications and drugs may alter the levels of a biomarker of interest (14).

When the database was searched for the test queries of copper, organophosphates, and particulate matter, only copper produced results (see Table 2). Some useful information from the

copper query results included how copper is processed by the body. It is thought that copper is reduced to $Cu1^{+}$ before it is absorbed by the gut via copper transport proteins. Once secreted into the bloodstream, copper remains bound to serum proteins until it is eliminated through bile. DrugBank also highlighted the role of copper in redox reactions where it is important for superoxide activity helping mitigate the presence of reactive oxygen species. Too much copper leads to cell damage due to the production of reactive oxygen species causing cell damage, which is thought to be its mechanism of toxicity. On the other hand, copper deficiency is also detrimental, causing connective tissue and neurological impairment. Interestingly, copper ions also affect fertility which is why some intrauterine devices contain copper.

Overall, this database is useful to research the effects of known medications and supplements taken by study subjects. However, data provided from the database was incomplete as tissue targets, half-life, and measurement details were not available, which are important for estimating exposure. Because this database is not tailored specifically to exposure, it has limited use for biomarker application.

### 3.2.3. PubChem

PubChem is a freely available database supported by the NIH Library of Medicine (15). PubChem compiles data from over 754 data sources including government agencies, publications, and chemical vendors. One unique feature of PubChem is the mechanism for outside, unaffiliated labs, investigators, and companies to submit data to be included in the database. These submissions must undergo a review process before being included. Modifications to database records are tracked and cataloged by date allowing the user to see when the last update occurred. Query results also include a detailed literature list. However, the majority of the information presented by PubChem are links to other sites, which while helpful, limits the utility of the database. If the user wishes to search chemicals, specifically, PubChem provides details on chemical properties including weight, melting point, solubility, and physical properties including color and odor. With respect to pharmacology and mechanisms of action, PubChem links the user back to DrugBank.

All three search queries provided results in PubChem. Both organophosphates and particulate matter (PM2.5 and PM10) provided results that linked to the CTD, which will be discussed in detail below. On the other hand, PubChem provided a detailed analysis of the chemical properties of copper along with a summary of toxicological effects, which are important criteria for understanding how copper exposure effects the body. PubChem also provided a table of copper associated disorders and diseases; however, this information also came from CTD. Finally, there is very little, if any, biomonitoring information, specifically, biofluids where copper can be measured and methods of detection. These results highlight the fact that PubChem is not the best source for exposure related searches and that specific chemical searches are more likely to be successful. On the other hand, PubChem does provide links to other databases, which can be helpful if the user is not sure where to start in their search.

### 3.2.4. Comparative Toxicogenomics Database (CTD)

The CTD is a freely available environmental exposure database supported by the National Institute of Environmental Health Studies (16). It is hosted by North Carolina State

University and updated monthly. The goal of the database is to further understand how environmental exposures affect human health. To do this, the database includes information linking chemicals to gene/protein interactions and disease development. The chemical-disease and gene-disease links include both curated and inferred associations. Curated disease associations are established by both the CTD and the OMIM database. Inferred associations are established by assuming that if a chemical has a known interaction with a gene and the gene has a known association with a disease that the chemical is associated with the disease. This comprehensive database also includes detailed references.

In the CTD, the user is able to perform a general search or specifically search by chemical, disease, gene, organism, pathway, and anatomy. The results summary page provides a list of the top interacting genes and distributes information in a series of tabs (Figure 5).

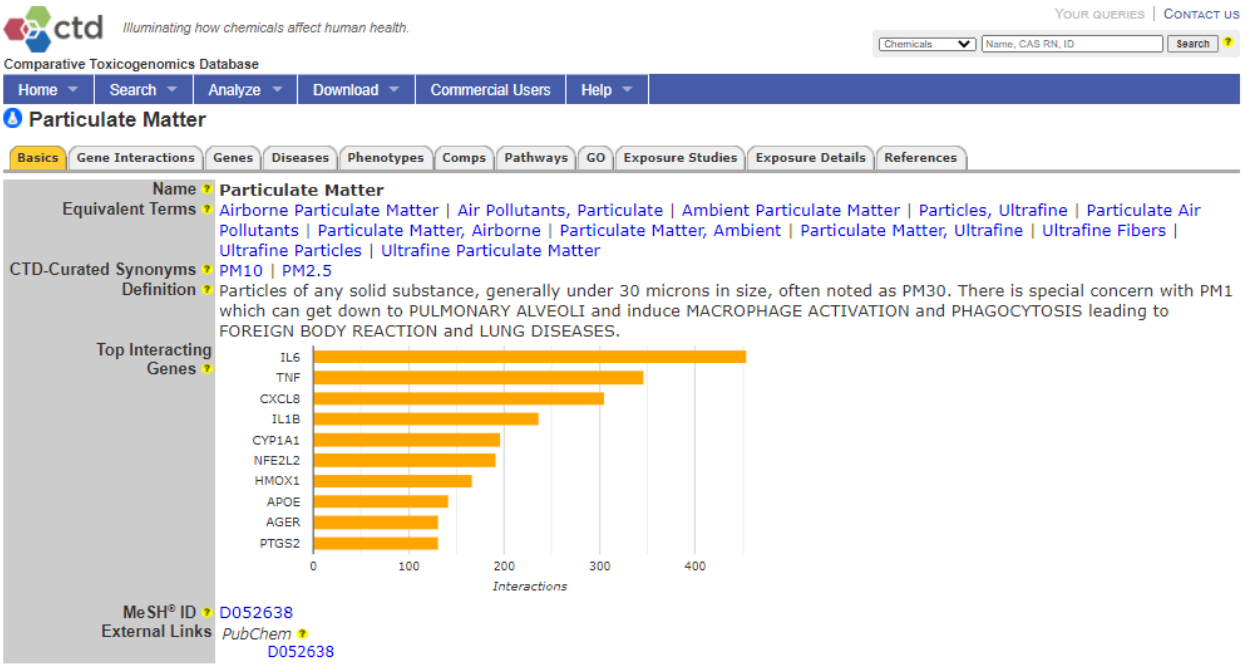
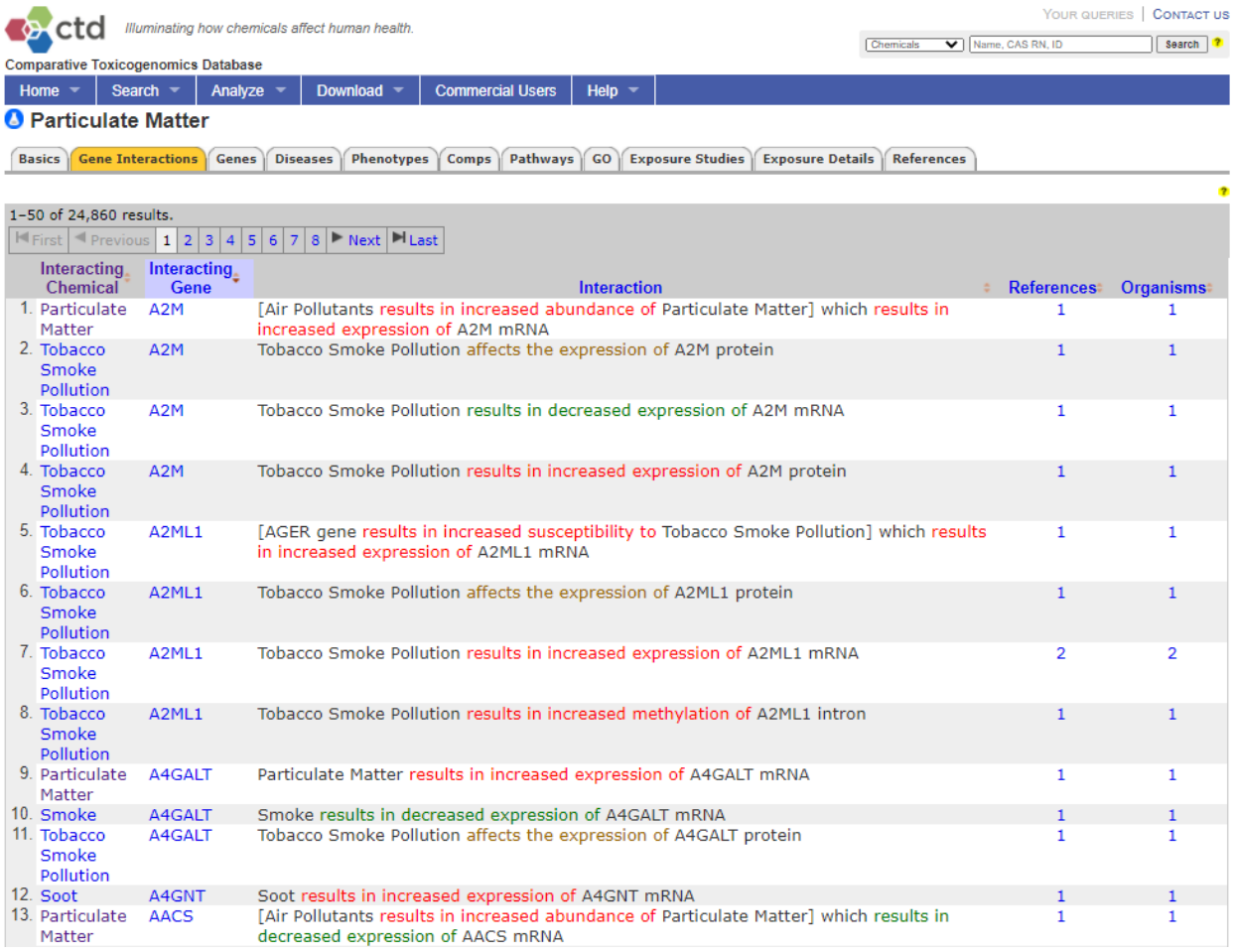



**Figure 5. Representative results page from CTD query of "particulate matter" (16)**

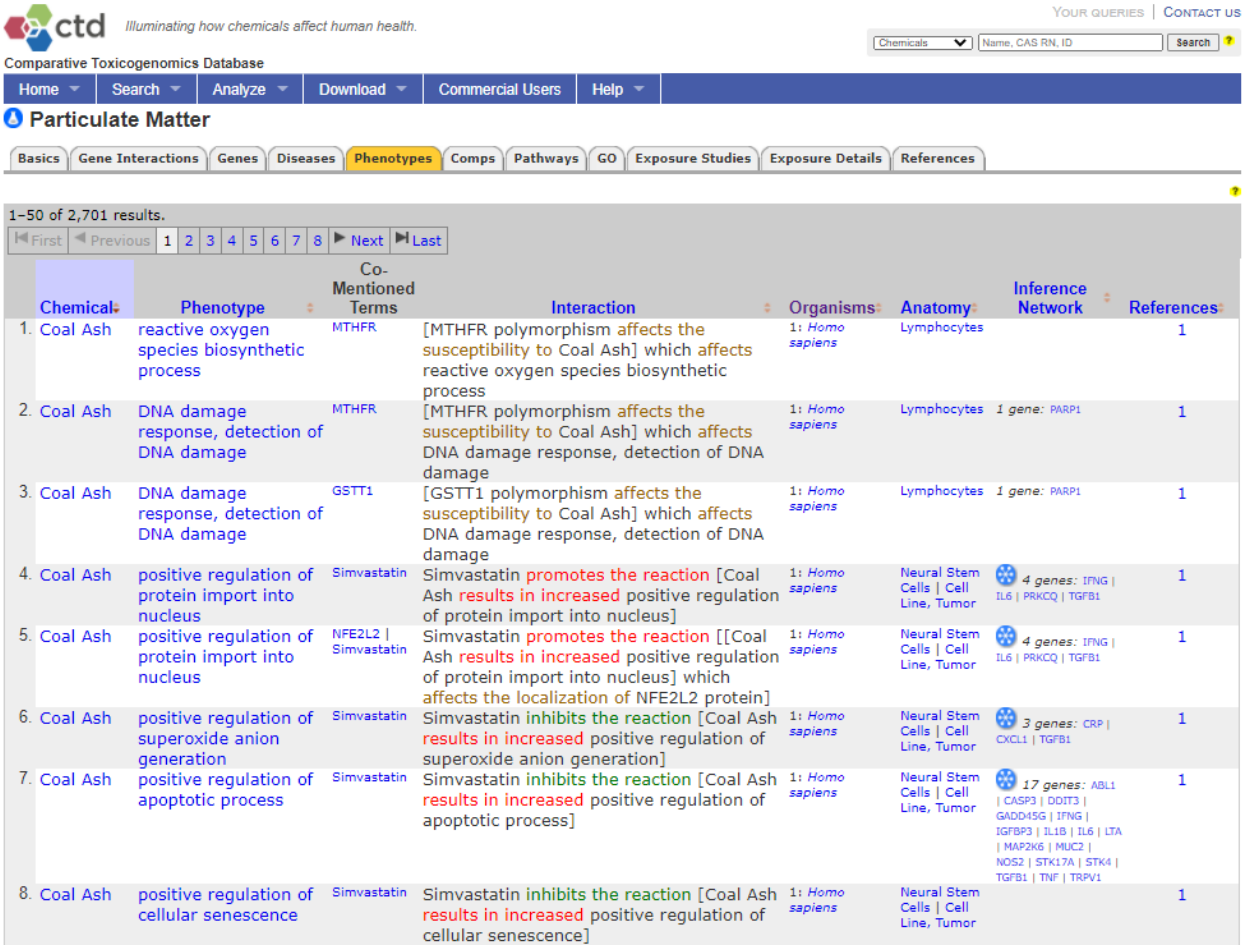
The "gene interactions" tab reveals a table of genes that the search term is documented to interact with along with whether the gene is increased or decreased due to the interacting chemical (Figure 6). These data are useful for identifying potential genetic markers of exposure.




| | Interacting Chemical | Interacting Gene | Interaction | References | Organisms |
|---|---|---|---|---|---|
| 1. | Particulate Matter | A2M | [Air Pollutants results in increased abundance of Particulate Matter] which results in increased expression of A2M mRNA | 1 | 1 |
| 2. | Tobacco Smoke Pollution | A2M | Tobacco Smoke Pollution affects the expression of A2M protein | 1 | 1 |
| 3. | Tobacco Smoke Pollution | A2M | Tobacco Smoke Pollution results in decreased expression of A2M mRNA | 1 | 1 |
| 4. | Tobacco Smoke Pollution | A2M | Tobacco Smoke Pollution results in increased expression of A2M protein | 1 | 1 |
| 5. | Tobacco Smoke Pollution | A2ML1 | [AGER gene results in increased susceptibility to Tobacco Smoke Pollution] which results in increased expression of A2ML1 mRNA | 1 | 1 |
| 6. | Tobacco Smoke Pollution | A2ML1 | Tobacco Smoke Pollution affects the expression of A2ML1 protein | 1 | 1 |
| 7. | Tobacco Smoke Pollution | A2ML1 | Tobacco Smoke Pollution results in increased expression of A2ML1 mRNA | 2 | 2 |
| 8. | Tobacco Smoke Pollution | A2ML1 | Tobacco Smoke Pollution results in increased methylation of A2ML1 intron | 1 | 1 |
| 9. | Particulate Matter | A4GALT | Particulate Matter results in increased expression of A4GALT mRNA | 1 | 1 |
| 10. | Smoke | A4GALT | Smoke results in decreased expression of A4GALT mRNA | 1 | 1 |
| 11. | Tobacco Smoke Pollution | A4GALT | Tobacco Smoke Pollution affects the expression of A4GALT protein | 1 | 1 |
| 12. | Soot | A4GNT | Soot results in increased expression of A4GNT mRNA | 1 | 1 |
| 13. | Particulate Matter | AACS | [Air Pollutants results in increased abundance of Particulate Matter] which results in decreased expression of AACS mRNA | 1 | 1 |

**Figure 6. Representative table from "Gene Interactions" tab of CTD results from search term "particulate matter" (16)**

The "Phenotypes" tab produces a table of phenotypes associated with the chemicals listed in the chemical column (Figure 7). This table includes targeted anatomy as well as the species associated with the phenotype making it easy to look for human related affects.



ctd Illuminating how chemicals affect human health.

Comparative Toxicogenomics Database

Your queries | Contact us

Chemicals | Name, CAS RN, ID | Search

Home | Search | Analyze | Download | Commercial Users | Help

Particulate Matter

Basics | Gene Interactions | Genes | Diseases | Phenotypes | Comps | Pathways | GO | Exposure Studies | Exposure Details | References

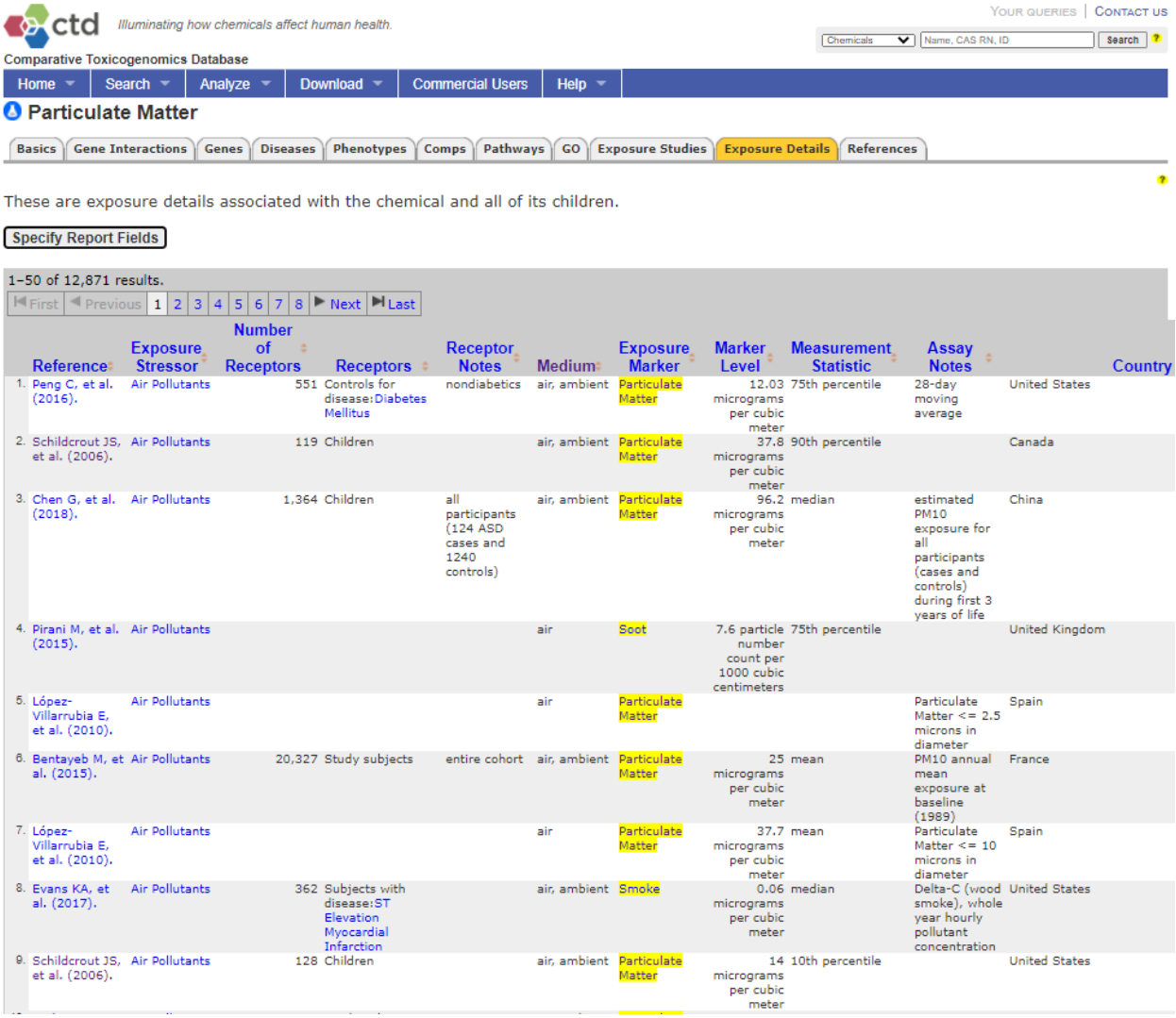
1–50 of 2,701 results.

First | Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Next | Last

| | Chemical | Phenotype | Co-Mentioned Terms | Interaction | Organisms | Anatomy | Inference Network | References |
|---|---|---|---|---|---|---|---|---|
| 1. | Coal Ash | reactive oxygen species biosynthetic process | MTHFR | [MTHFR polymorphism affects the susceptibility to Coal Ash] which affects reactive oxygen species biosynthetic process | 1: *Homo sapiens* | Lymphocytes | | 1 |
| 2. | Coal Ash | DNA damage response, detection of DNA damage | MTHFR | [MTHFR polymorphism affects the susceptibility to Coal Ash] which affects DNA damage response, detection of DNA damage | 1: *Homo sapiens* | Lymphocytes | *1 gene:* PARP1 | 1 |
| 3. | Coal Ash | DNA damage response, detection of DNA damage | GSTT1 | [GSTT1 polymorphism affects the susceptibility to Coal Ash] which affects DNA damage response, detection of DNA damage | 1: *Homo sapiens* | Lymphocytes | *1 gene:* PARP1 | 1 |
| 4. | Coal Ash | positive regulation of protein import into nucleus | Simvastatin | Simvastatin promotes the reaction [Coal Ash results in increased positive regulation of protein import into nucleus] | 1: *Homo sapiens* | Neural Stem Cells \| Cell Line, Tumor | *4 genes:* IFNG \| IL6 \| PRKCQ \| TGFB1 | 1 |
| 5. | Coal Ash | positive regulation of protein import into nucleus | NFE2L2 \| Simvastatin | Simvastatin promotes the reaction [[Coal Ash results in increased positive regulation of protein import into nucleus] which affects the localization of NFE2L2 protein] | 1: *Homo sapiens* | Neural Stem Cells \| Cell Line, Tumor | *4 genes:* IFNG \| IL6 \| PRKCQ \| TGFB1 | 1 |
| 6. | Coal Ash | positive regulation of superoxide anion generation | Simvastatin | Simvastatin inhibits the reaction [Coal Ash results in increased positive regulation of superoxide anion generation] | 1: *Homo sapiens* | Neural Stem Cells \| Cell Line, Tumor | *3 genes:* CRP \| CXCL1 \| TGFB1 | 1 |
| 7. | Coal Ash | positive regulation of apoptotic process | Simvastatin | Simvastatin inhibits the reaction [Coal Ash results in increased positive regulation of apoptotic process] | 1: *Homo sapiens* | Neural Stem Cells \| Cell Line, Tumor | *17 genes:* ABL1 \| CASP3 \| DDIT3 \| GADD45G \| IFNG \| IGFBP3 \| IL1B \| IL6 \| LTA \| MAP2K6 \| MUC2 \| NOS2 \| STK17A \| STK4 \| TGFB1 \| TNF \| TRPV1 | 1 |
| 8. | Coal Ash | positive regulation of cellular senescence | Simvastatin | Simvastatin inhibits the reaction [Coal Ash results in increased positive regulation of cellular senescence] | 1: *Homo sapiens* | Neural Stem Cells \| Cell Line, Tumor | | 1 |

**Figure 7. Representative table under "Phenotype" tab of CTD results from search term "particulate matter" (16)**

Finally, under the "Exposure Details" tab, a table is generated that breaks down each reference by exposure stressor, receptors (or study population), measurement medium, exposure marker, and disease implications (Figure 8). These results can be expanded and narrowed based on the check boxes above the table. All of the tables generated by CTD allow the user to sort the data in each column in order to focus on a specific area of interest. Unfortunately, sorting on the website can be tedious so CTD also allows the user to export all tables to excel which makes sorting data and finding key terms of interest much easier.



ctd Illuminating how chemicals affect human health.

Comparative Toxicogenomics Database

Your queries | Contact us

Chemicals | Name, CAS RN, ID | Search

Home | Search | Analyze | Download | Commercial Users | Help

Particulate Matter

Basics | Gene Interactions | Genes | Diseases | Phenotypes | Comps | Pathways | GO | Exposure Studies | Exposure Details | References

These are exposure details associated with the chemical and all of its children.

Specify Report Fields

1–50 of 12,871 results.

First | Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Next | Last

| | Reference | Exposure Stressor | Number of Receptors | Receptors | Receptor Notes | Medium | Exposure Marker | Marker Level | Measurement Statistic | Assay Notes | Country |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | Peng C, et al. (2016). | Air Pollutants | 551 | Controls for disease:Diabetes Mellitus | nondiabetics | air, ambient | Particulate Matter | 12.03 micrograms per cubic meter | 75th percentile | 28-day moving average | United States |
| 2. | Schildcrout JS, et al. (2006). | Air Pollutants | 119 | Children | | air, ambient | Particulate Matter | 37.8 micrograms per cubic meter | 90th percentile | | Canada |
| 3. | Chen G, et al. (2018). | Air Pollutants | 1,364 | Children | all participants (124 ASD cases and 1240 controls) | air, ambient | Particulate Matter | 96.2 micrograms per cubic meter | median | estimated PM10 exposure for all participants (cases and controls) during first 3 years of life | China |
| 4. | Pirani M, et al. (2015). | Air Pollutants | | | | air | Soot | 7.6 particle number count per 1000 cubic centimeters | 75th percentile | | United Kingdom |
| 5. | López-Villarrubia E, et al. (2010). | Air Pollutants | | | | air | Particulate Matter | | | Particulate Matter <= 2.5 microns in diameter | Spain |
| 6. | Bentayeb M, et al. (2015). | Air Pollutants | 20,327 | Study subjects | entire cohort | air, ambient | Particulate Matter | 25 micrograms per cubic meter | mean | PM10 annual mean exposure at baseline (1989) | France |
| 7. | López-Villarrubia E, et al. (2010). | Air Pollutants | | | | air | Particulate Matter | 37.7 micrograms per cubic meter | mean | Particulate Matter <= 10 microns in diameter | Spain |
| 8. | Evans KA, et al. (2017). | Air Pollutants | 362 | Subjects with disease:ST Elevation Myocardial Infarction | | air, ambient | Smoke | 0.06 micrograms per cubic meter | median | Delta-C (wood smoke), whole year hourly pollutant concentration | United States |
| 9. | Schildcrout JS, et al. (2006). | Air Pollutants | 128 | Children | | air, ambient | Particulate Matter | 14 micrograms per cubic meter | 10th percentile | | United States |

**Figure 8. Representative table under the "Exposure Details" tab of CTD results from the search term "particulate matter" (16)**

One feature that CTD offers is the ability to perform analyses within the database comparing multiple genes, phenotypes, chemicals, diseases, and pathways. There are five analysis tools provided by the database (Figure 9).

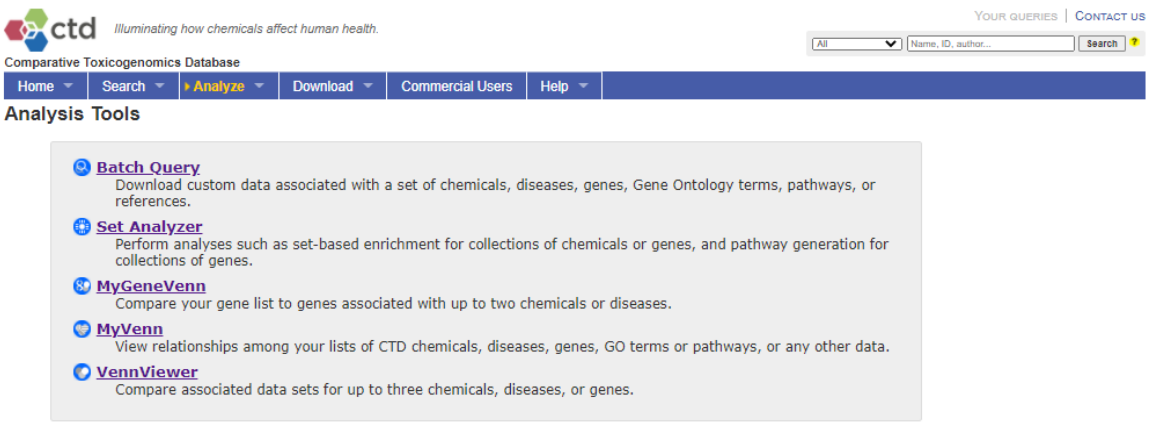



**Figure 9. Analysis tools available in CTD (16)**

Of particular interest is the VennViewer where the user can identify common genes, phenotypes, diseases and pathways that are in common and separate from several different queries of interest (chemicals, genes, phenotypes, diseases). For example, when the chemical inquiries "PM2.5, copper, and carbon monoxide" are used, Venn diagrams can be generated showing overlapping and independent diseases, pathways, phenotypes and gene ontology terms (Figure 10). This is especially useful when investigating exposures with complex mixtures. Each section of the diagram links to a list of pathways, diseases, phenotypes, or related gene ontology terms, which are then linked to information pages about the selected item. This provides a targeted investigation list for both individual exposures and multiple exposures at the same time. Finally, these Venn diagrams can also be generated using gene expression changes. The data can be grouped by any change, genes that are increased and genes that are decreased. This information is helpful for identifying genetic changes associated with exposure. As mentioned earlier, specific information pages are linked and provide valuable data and resources. In this instance, carbon monoxide was used as a related exposure of interest to demonstrate the utility of VennViewer in addition to PM2.5 and copper. Of the six curated diseases associated with PM2.5, copper, and carbon monoxide, heart failure is listed and links to a CTD with all of the tabs described above.

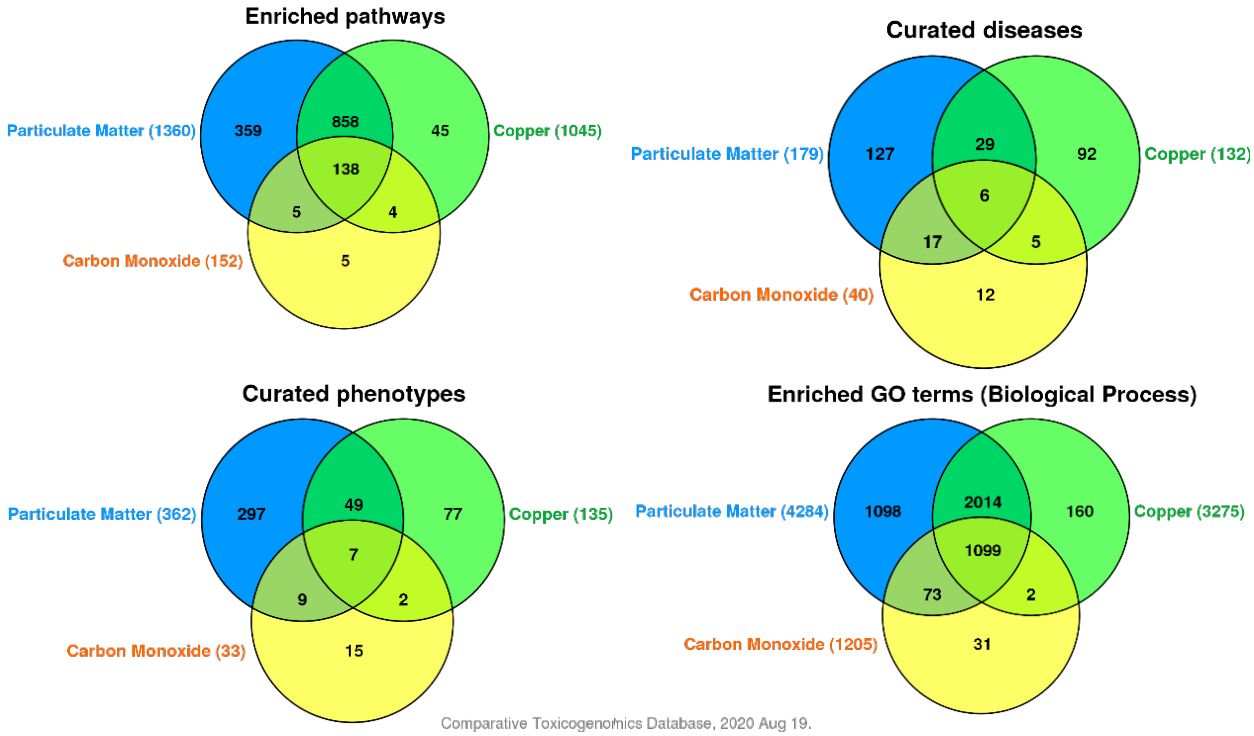



**Figure 10. Representative Venn diagrams utilizing the CTD database**

Overall, the CTD is quite useful for biomarker research because of the comprehensive nature. Not only can the user identify diseases associated with exposures of interest, but also genes, markers, and phenotypes to be aware of. This database is a particularly good resource for narrowing down potential genetic and biological markers of exposure and resulting disease development.

### 3.2.5. Chemical Entities of Biological Interest

Chemical Entities of Biological Interest (ChEBI) is a freely available database of small chemical compounds both naturally produced and synthetic (17). ChEBI compiles data from several sources including the integrated relational enzyme database, ChEMBL, and the Protein Data Bank in Europe (PDBeChem). ChEBI provides chemical information for search terms including formula, mass, registry numbers, and references, if any, to chemical-gene and chemical-protein interactions. Unfortunately, this database is *not* designed to include genes, nucleic acids, proteins, and/or peptides, limiting the utility of this database for biomarker related inquiries. Because the database is chemical specific, the search query "particulate matter" did not provide any results. The search term "organophosphate AND pesticide" revealed 30 resulting pages, including one for "organophosphate insecticides" while "copper" had the most results (63 entries) corresponding to different copper containing chemical compounds. These results pages mainly contained chemical information and no data on exposure or disease implications.

### 3.2.6. Online Mendelian Inheritance in Man (OMIM)

OMIM is a publicly available, continuously updated database curated by the McKusick-Nathans Institute of Genetic Medicine and The Johns Hopkins University School of Medicine (18). Data included in OMIM is gathered from peer-reviewed biomedical publications focused on the relationship between genotype variation and phenotype. Specifically, OMIM catalogs genes and gene mutations that are associated, either directly or indirectly, with a phenotype or disease. A search query includes a list of gene or gene mutation hits that include the search terms. The result hit links to a page with a detailed literature review of the gene/phenotype relationship from basic science studies to clinical trials. For example, when the term "organophosphate poisoning" was searched, the only hit was for Paraoxonase1 (PON1). Under the results page, a PON1 autosomal dominant mutation was identified as conferring susceptibility to organophosphate poisoning. The review includes gene structure, function, and mapping, a molecular genetics overview, genotype/phenotype analysis, and animal models. This database may be useful for identifying biomarkers of susceptibility. In fact, the OMIM results for "organophosphate poisoning" indicate that changes in plasma PON1 are associated with phenotype development. However, with respect to exposure, this database is limited to those genes and gene mutations that are associated with a phenotype not necessarily due to an exposure.

### 3.2.7. Exposome-explorer

Exposome-explorer is a database developed by the International Agency for Research on Cancer in collaboration with the University of Alberta in Canada released in 2016 and last updated in September of 2019 (19, 20). This database is freely available and provides an analysis of known biomarkers of exposure to dietary factors and environmental contaminants from published literature. Search query results include general tabs for chemical data, publications, biomarker data, and exposure data. Chemical data contains compound structure, Chemical Abstracts Service (CAS) number, description, and molecular weight. Importantly, the biomarker data section includes a summary table from published studies of biospecimens, biomarkers, and concentrations along with a link to the reference. Additionally, associations with exposure and cancer are also listed. Finally, the section on exposure data displays known correlations with the exposure in specific populations and associations of the exposure with biomarkers. Unfortunately, this database is not comprehensive. According to database statistics, only 811 publications have been evaluated identifying 907 biomarkers. Based solely on the "biomarker AND military" PubMed search, over 6,400 results were displayed, indicating that this database only scratches the surface of published literature.

There are a few benefits of the Exposome-Explorer database. Under the biomarker data tab is a list of all of the biomarkers in the database and the number of publications associated with the biomarker. The user can also look at biomarkers correlated with exposures where the biospecimen, correlation type, intake (or exposure), and population are listed. Finally, under the classifications tab, the user can select specific categories including biospecimens, analytical methods, and cohorts. For example, under biospecimens category, a list is generated showing the specific biospecimen (e.g. blood, urine, plasma, hair, nails.) along with the number of biospecimen subtypes or collection time. The biospecimen name is linked to a table of the biomarkers measurable in the biospecimen of interest, the concentration, normalization method,

and publication. It may be useful to search biospecimen categories to investigate whether the biomarker of interested can be measured in the specific biospecimen. However, the utility for military exposure biomarkers is limited.

### 3.2.8. CDC National Biomonitoring Program

The CDC National Biomonitoring Program provides an assessment of exposures to the U.S. population. The website provides links to a list of 25 chemical factsheets and 93 biomonitoring summaries. The biomonitoring summary contains a brief literature review on general chemical information, exposure routes, physiological effects, metabolism of chemicals, and biomonitoring applications including biofluids used for exposure evaluation and references. Unfortunately, the list of summaries is limited, making this database useful if the user knows what compound to target and if the compound happens to be cataloged in the program. Of the test queries utilized in this review, the biomonitoring program only had a chemical factsheet for organophosphorus insecticides, specifically dialkyl phosphate. Included in the factsheet are common exposure routes, health effects, and links to the EPA and NIOSH. The biomonitoring summary provided a more detailed literature review of dialkyl phosphate, including a list of metabolites of exposure that can be measured in urine. Overall, this database has limited utility for biomarker application work. This is especially true given the fact that the page has not been reviewed since April 7, 2017 and the limited list of chemicals included.

### 3.2.9. International Peer Reviewed Chemical Safety Information (INCHEM)

INCHEM is an international initiative to consolidate current, internationally peer-reviewed chemical safety-related publications and database records from international bodies, for public access. In collaboration with the WHO and the Canadian Center for Occupational Health and Safety (CCOHS), INCHEM provides rapid electronic access to full-text documents on chemical risks and safety. Search results provide a list of full-text reports associated with the search term (Figure 11). Reports contain a comprehensive review of chemical information including chemical properties, analytical methods, sources of human environmental exposure, effects of exposure in animals and humans, and evaluation of health risks. Importantly, INCHEM reports contain information on routes of exposure, summaries of toxicity, and effects of repeated and long-term exposure. Because the results are reports, the information can be difficult to sort through to find items of interest.

Additionally, the reports are not necessarily specific to the search term. When “PM2.5” was used as the search term, reports on diesel exhaust, copper, and manganese were some of the relevant results (Figure 11). These results make sense, as these exposures can be present in particulate matter; however, a summary report dedicated to particulate matter in general was not found. Additionally, while these reports are relatively comprehensive with respect to chemical properties, they do not provide information on gene and protein targets of exposures, which are important considerations when investigating biomarkers of exposure.

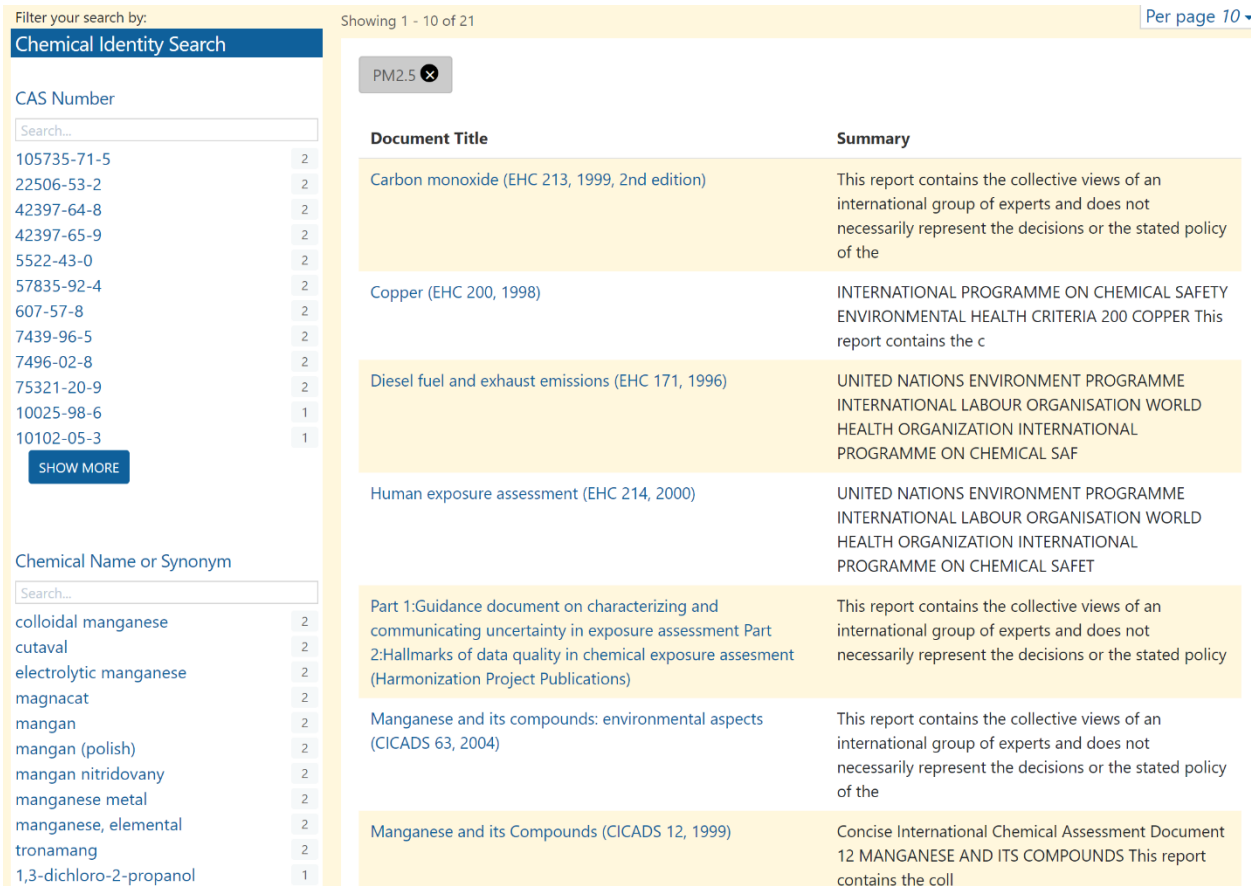



**Figure 11. Representative INCHEM search results**

### 3.2.10. Measuring Environmental Biomarkers-Laboratories (MEB-Labs Database)

The MEB-Labs database is publicly available developed by the Integrative Health Sciences Core at the University of Cincinnati. The database contains information on measurable chemical biomarkers, analytical techniques, limits of detection and references. The user is able to search by analyte (13 classes), a list of chemicals, and a list of laboratories. The database is limited in scope because only a fraction of literature (96 publications) and chemicals/biomarkers (160 total) are included. Search queries provide a table of related biomarkers, labs where the analysis is performed, the sample used for measurement, and the limits of detection (Figure 12). This database is useful to identify biofluids that may be used for detection as well as the labs that perform the work so follow up research may be completed to identify methods of detection. Unfortunately, due to the limited list of analytes and chemicals, there are substantial gaps in the database. For example, when using the analyte class "metals" to generate a list of results, copper is not included in the list. Unfortunately, during final editing of this report, it appears that this database has been suspended. Attempts to select a search category resulted in a server error.

ANALYTES Excel Plain Text Print Start Over Refine Query

Jump to: Summary Analytes

| Name of lab | Class of analytes | Name of analyte | Analyze | Sample source | Limits of detection |
|---|---|---|---|---|---|
| CDC Environmental Lab | Organophosphate Pesticides | 2-(Diethylamino)-6-methylpyrimidin-4-ol/one | Yes | Urine | Urine: 0.2 |
| CDC Environmental Lab | Organophosphate Pesticides | 2-Isopropyl-4-methyl-6-hydroxypyrimidine | Yes | Urine | Urine: 0.7 |
| CDC Environmental Lab | Organophosphate Pesticides | 3,5,6-Trichloro-2-pyridinol | Yes | Urine | Urine: 0.4 |
| CDC Environmental Lab | Organophosphate Pesticides | 3-Chloro-7-hydroxy-4-methyl-2H-chromen-2-one/ol | Yes | Urine | Urine: 0.2 |
| CDC Environmental Lab | Organophosphate Pesticides | Diethyldithiophosphate | Yes | Urine | Urine: 0.1 |
| CDC Environmental Lab | Organophosphate Pesticides | Diethylphosphate | Yes | Urine | Urine: 0.2 |
| CDC Environmental Lab | Organophosphate Pesticides | Diethylthiophosphate | Yes | Urine | Urine: 0.1 |
| CDC Environmental Lab | Organophosphate Pesticides | Dimethyldithiophosphate | Yes | Urine | Urine: 0.1 |
| CDC Environmental Lab | Organophosphate Pesticides | Dimethylphosphate | Yes | Urine | Urine: 0.5 |
| CDC Environmental Lab | Organophosphate Pesticides | Dimethylthiophosphate | Yes | Urine | Urine: 0.4 |
| CDC Environmental Lab | Organophosphate Pesticides | Malathion Dicarboxylic Acid | No | | |
| CDC Environmental Lab | Organophosphate Pesticides | Para-Nitrophenol | Yes | Urine | Urine: 0.1 |
| National Exposure Research Laboratory | Organophosphate Pesticides | 2-(Diethylamino)-6-methylpyrimidin-4-ol/one | No | | |
| National Exposure Research Laboratory | Organophosphate Pesticides | 2-Isopropyl-4-methyl-6-hydroxypyrimidine | No | | |
| National Exposure Research Laboratory | Organophosphate Pesticides | 3,5,6-Trichloro-2-pyridinol | Yes | Urine | Urine: 10ppb by ELISA in urine, food, dust and soil |
| National Exposure Research Laboratory | Organophosphate Pesticides | 3-Chloro-7-hydroxy-4-methyl-2H-chromen-2-one/ol | No | | |
| National Exposure Research Laboratory | Organophosphate Pesticides | Diethyldithiophosphate | No | | |
| National Exposure Research Laboratory | Organophosphate Pesticides | Diethylphosphate | No | | |
| National Exposure Research Laboratory | Organophosphate Pesticides | Diethylthiophosphate | No | | |

**Figure 12. List of results from MEB-Labs when searching the analyte class "organophosphate pesticides"**

## 3.3 Department of Defense Databases

In addition to open access databases, there are also several databases that are specific to the DoD. These databases are commonly used to document epidemiological and health data from military members. These databases are not publicly available and some require rigorous steps to obtain access. A list of the databases with access requirements are shown in (Table 3).

**Table 3. Information Related to DoD Databases and Studies**

| Database/Study | Website | Access |
|---|---|---|
| DOEHRS | https://doehrs-ih.csd.disa.mil/ | Requested through website |
| MCS | https://www.millenniumcohort.org/about | 3-5 page outline, then 10-15 page proposal |
| AFHS | ISBN 978-0-309-33914-8 | PDF of book (2015) freely accessible. To access data – Letter of Intent, then proposal |
| DoD-CCR | https://www.jpc.capmed.mil/education/dodccrs/index.asp | unknown |
| ESSENCE | https://health.mil/Military-Health-Topics/Combat-Support/Armed-Forces-Health-Surveillance-Branch/Integrated-Biosurveillance/ESSENCE | unknown |
| ASIMS | https://www.health.mil/Reference-Center/Fact-Sheets/2020/09/02/ASIMS | CAC authentication required. Must request account. |

CAC – Common Access Card

A summary of DoD database purpose and content is shown in Table 4. Descriptions are also included in the following sections.

**Table 4. Summary of DoD Databases and Studies**

| Database Name | Purpose | Variables/Content |
|---|---|---|
| Defense Occupational and Environmental Health Reporting System (DOEHRS) (21) | To manage documentation of threats with completed pathways of exposure; unclassified | Industrial hygiene assessment, acceptable level (Y/N), start date, exposure rating, health effect rating, hazard name |
| Air Force Health Study (Ranch Hand) (AFHS) (23) | Compare health effects in Vietnam veterans exposed or not to herbicides (Agent Orange) | Personnel/partners/kids medical records, physical exams (medical tests, clinical endpoints), serum samples, education, income, fertility/marital status, lifestyle behaviors, toxic substances exposure, herbicide exposure, military experience |
| Millennium Cohort Study (MCS) | Prospective epidemiological research post-Gulf War to evaluate the impact of military exposures, including deployment, on long-term health outcomes | General health, health behaviors, physical symptoms, diagnosis, pregnancy/infertility, out/in patient visits, alternative medicine/other facility use, vaccines, potential exposures (combat, injuries, burn pit, chemical/pesticide/uranium, assault, etc.), military life, employment, education, homelessness, anger, stress events |
| Automated Central Tumor Registry (ACTUR) / DoD Consolidated Cancer Registry System (CCR) | Current cancer registry system in use in the Department of Defense to track cancer and tumor development; part of the Defense Enrollment Eligibility System (DEERS) | name, Social Security Number, phone numbers, age, addresses, emails, medical record numbers, race, and date of birth are collected; collected about DoD Service members (active or retired) and their dependents; oncology data (specifics unknown) |
| Electronic Surveillance System for the Early Notification of Community-based Epidemics (ESSENCE) (25) | Monitor worldwide DoD outpatient health care data for certain medical syndromes to detect trends of public health interest; Allow DHA epidemiologists and public health officers to obtain medical situational awareness and investigate reportable disease events | Patient visit data: arrival time, facility, illness category, chief complaint, triage notes, tests ordered, age, zip code, sex, medical record#, system identification, International Classification of Disease (ICD)-10 codes |
| Aeromedical Services Information Management System (ASIMS) (26) | Installation wide system for unit commanders to access their Airmen's medical readiness requirements (AFI 10-250). Allow data sharing between personnel and medical readiness information systems. | Immunization records, health assessments (baseline, pre deployment, post deployment), medical clearances, occupation health, fitness limitations, pre and post deployment medical processing. Interfaces with 19 systems. |

DoD – Department of Defense, DHA – Defense Health Agency, ICD – International Classification of Disease. AFI – Air Force Instruction

### 3.3.1. Defense Occupational and Environmental Health Reporting System (DOEHRS)

The DOEHRS is a product of the Defense Health Agency (DHA) and is managed by the Defense Health Service System (DHSS) (Table 4). According to the DOEHRS Factsheet, this database is designed to be an individual, longitudinal exposure record for military members and includes workplace practices, environmental surveillance and public health information, and environmental exposure registries (21). Some of the information tracked by the database include biological, chemical and physical health hazards. The database is updated regularly. However, it does not link exposures with measurable biomarkers. Biomarker and biomonitoring information would be a useful update for this database.

### 3.3.2. Air Force Health Study (AFHS)

The AFHS is a longitudinal investigation (1982 – 2006) following a cohort of military personnel exposed to the herbicide, Agent Orange, in southern Vietnam (Table 4) (22). While not a database, this study used questionnaires to document the long-term health consequences of exposure to Agent Orange including medical history, marital and fertility information, toxic substances exposure, and other lifestyle behaviors. Biospecimens, including urine, semen, and blood, were collected over 6 study visits during the period of performance. These samples are now banked and can be utilized for further studies on biomarkers of exposure and effect (23).

### 3.3.3. Millennium Cohort Study (MCS)

The MCS is a prospective study designed to follow a large group of military members over a long period of time (Table 4). Enrollment in this study began in 2001 and has continued up through 2019 with over 200,000 participants enrolled. The study utilizes surveys to collect a wide array of data including medical history, in/out patient visits, deployments, education, and potential exposures. This study targets military personnel that may have served in Iraq and Afghanistan (24). The data collected in this study can help link exposures with long-term health effects helping identify biomarkers of effect that may be monitored. Some of the targeted long-term effects include chronic pain, obesity, and reproductive health.

### 3.3.4. DoD Consolidated Cancer Registry System (CCR)

The DoD CCR was created through the Deputy Assistant Director for Medical Affairs to collect and document data regarding tumor and cancer development in military members (Table 4). The DoD CCR works in collaboration with the Automated Central Tumor Registry (ACTUR) to monitor and track cancer data for all DoD beneficiaries. If this registry also contains information regarding potential exposures experienced while active in the military, then cancer development may be linked with exposure and thus biomarkers could be utilized to monitor cancer risk and progression.

### 3.3.5. Electronic Surveillance System for Early Notification of Community-based Epidemics (ESSENCE)

Developed in 2004, this program is designed to increase Military Health System (MHS) monitoring capability for detection of health threats that may impact force health readiness (25)

(Table 4). This system utilizes the same software as the CDC allowing collaboration with civilian counterparts during outbreaks. ESSENCE monitors the direct care of the MHS population by analyzing DoD healthcare data continuously. The database also allows queries by disease, reported medical event, and injury categories. Important for biomarker discovery, this database provides longitudinal health monitoring allowing for the identification of potential chronic health effects.

### 3.3.6. Aeromedical Services Information Management System (ASIMS)

ASIMS is a web-based system utilized by Active Duty AF to track medical readiness and health data including immunizations, health assessments, both pre and post deployment, and medical clearance records (26) (Table 4). This system tracks National Defense Authorization Act and DoD mandated health requirements for AF members along with occupational health assessments. Included in the records are fitness, duty, and mobility restrictions, which can be important confounding factors for biomarker detection. Importantly, ASIMS interfaces with 19 other databases simplifying data retrieval. Furthermore, health information collected before and after deployment can help pinpoint occupational exposures that may have occurred during deployment.

## 4.0 CONCLUSION

The goal of this report was to review open access and DoD databases to determine their utility for informing biological measurements during exposure assessments. Ten open access databases were evaluated and tested for utility using three military exposures of interest: copper, particulate matter, and organophosphates. Most of the databases did not contain extensive data on military-relevant exposures. The CTR provided the most comprehensive and useful results by documenting biological specimens collected, concentrations of chemicals and markers in samples, and citations for corresponding literature. The CTD is one of the only databases to allow comparisons of several different exposures to identify independent and overlapping genes, phenotypes, and diseases. Six Department of Defense databases were evaluated, and results found that these databases provide information related to individual warfighter health and exposure but do not adequately identify measurable biomarkers and methods for detection that can be applied in exposure assessments.

Based on our analysis, none of the databases evaluated can be used as a single reference to inform biomarker assessment for exposure assessments in the military. Therefore, either a new database is needed or existing databases need to be updated. An ideal database for biomarker assessment of exposure would include information about routes of exposure, relevant concentrations and durations of exposure, metabolites, biofluids where metabolites can be detected, half-life and stability of metabolites in each fluid, methods of detection, as well as references to studies where data was generated with details about the study, including population demographics and size. These data would be available for military-specific exposures, including jet fuel, firing range emissions, hypoxia, g-force, impulse noise, and exposures encountered during deployment. This database would automatically import information from open access databases and would be straightforward to populate with new data. Since much of the data would be drawn from military studies, it would ideal if it were secure with access extended to DoD employees and DoD contractors.

## REFERENCES


1. Martin, N. J., Richards, E. E., & Kirkpatrick, J. S. (2011). Exposure science in U.S. military operations: a review. *Mil Med, 176*(7 Suppl), 77-83. doi:10.7205/milmed-d-11-00098
2. Krahl, P. L., Benchoff, E., Go, Y. M., Jones, D. P., Smith, M. R., Walker, D. I., et al. (2019). Advances in Comprehensive Exposure Assessment: Opportunities for the US Military. *J Occup Environ Med, 61 Suppl 12*, S5-S14. doi:10.1097/JOM.0000000000001677
3. May, L. M., Weese, C., Ashley, D. L., Trump, D. H., Bowling, C. M., & Lee, A. P. (2004). The recommended role of exposure biomarkers for the surveillance of environmental and occupational chemical exposures in military deployments: policy considerations. *Mil Med, 169*(10), 761-767. doi:10.7205/milmed.169.10.761
4. Camilla A. Mauzy, L. M. K. B., Coleen P. Baird, and Carole, & Tinklepaugh, N. (2015). Considerations Regarding Biomonitoring. In C. P. Baird, D. K. Harkins, & U. h. b. g. c. b. i. L. T. Borden Institute (Eds.), *Airborne Hazards Related to Deployment*: Borden Institute, US Army Medical Department Center and School.
5. Califf, R. M. (2018). Biomarker definitions and their applications. *Exp Biol Med (Maywood), 243*(3), 213-221. doi:10.1177/1535370217750088
6. DeBord, D. G., Burgoon, L., Edwards, S. W., Haber, L. T., Kanitz, M. H., Kuempel, E., et al. (2015). Systems Biology and Biomarkers of Early Effects for Occupational Exposure Limit Setting. *J Occup Environ Hyg, 12 Suppl 1*, S41-54. doi:10.1080/15459624.2015.1060324
7. ACGIH. (2018). *2018 ACGIH TLVs and BEIs: Based on the Documentation of the Threshold Limit Values for Chemical Substances and Physical Agents & Biological Exposure Indices*. Cincinnati, OH: American Conference of Governmental Industrial Hygienists.
8. Chen, X. H., Huang, S., & Kerr, D. (2011). Biomarkers in clinical medicine. *IARC Sci Publ*(163), 303-322.
9. Mussali-galante, P., Tovar-Sanchez, E., Valverde, M., & Rojos Del Castillo, E. (2013). Biomarkers of exposure for assessing environmental metal pollution: From molecules to ecosystems. *Rev. Int. Contam. Ambie., 29*(1), 117-140.
10. NIOSH. (2010). *NIOSH Pocket Guide to Chemical Hazards*. (2010-168c). Cincinnati, OH Retrieved from https://www.cdc.gov/niosh/npg/.
11. Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vazquez-Fresno, R., et al. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res, 46*(D1), D608-D617. doi:10.1093/nar/gkx1089
12. Migliaccio, C. T., Bergauff, M. A., Palmer, C. P., Jessop, F., Noonan, C. W., & Ward, T. J. (2009). Urinary levoglucosan as a biomarker of wood smoke exposure: observations in a mouse model and in children. *Environ Health Perspect, 117*(1), 74-79. doi:10.1289/ehp.11378
13. Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., et al. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res, 46*(D1), D1074-D1082. doi:10.1093/nar/gkx1037
14. Mayeux, R. (2004). Biomarkers: potential uses and limitations. *NeuroRx, 1*(2), 182-188. doi:10.1602/neurorx.1.2.182

15. Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2019). PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res, 47*(D1), D1102-D1109. doi:10.1093/nar/gky1033
16. Davis, A. P., Grondin, C. J., Johnson, R. J., Sciaky, D., McMorran, R., Wiegers, J., et al. (2019). The Comparative Toxicogenomics Database: update 2019. *Nucleic Acids Res, 47*(D1), D948-D954. doi:10.1093/nar/gky868
17. Hastings, J., Owen, G., Dekker, A., Ennis, M., Kale, N., Muthukrishnan, V., et al. (2016). ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Res, 44*(D1), D1214-1219. doi:10.1093/nar/gkv1031
18. Eick-Cost, A. A., Hu, Z., Rohrbeck, P., & Clark, L. L. (2017). Neuropsychiatric Outcomes After Mefloquine Exposure Among U.S. Military Service Members. *Am J Trop Med Hyg, 96*(1), 159-166. doi:10.4269/ajtmh.16-0390
19. Neveu, V., Moussy, A., Rouaix, H., Wedekind, R., Pon, A., Knox, C., et al. (2017). Exposome-Explorer: a manually-curated database on biomarkers of exposure to dietary and environmental factors. *Nucleic Acids Res, 45*(D1), D979-D984. doi:10.1093/nar/gkw980
20. Neveu, V., Nicolas, G., Salek, R. M., Wishart, D. S., & Scalbert, A. (2020). Exposome-Explorer 2.0: an update incorporating candidate dietary biomarkers and dietary associations with cancer risk. *Nucleic Acids Res, 48*(D1), D908-D912. doi:10.1093/nar/gkz1009
21. DHA. (2019). *Defense Occupational and Environmental Health Readiness System – Industrial Hygiene (DOEHRS-IH) Fact Sheet*. Washington, DC Retrieved from https://health.mil/Reference-Center/Fact-Sheets/2019/04/05/DOEHRS-IH.
22. Styka, A. N., & Butler, D. A. (2015). The Air Force Health Study Data and Specimens as a Resource for Researchers. *Mil Med, 180*(10 Suppl), 79-84. doi:10.7205/MILMED-D-15-00001
23. . (2015). In *The Air Force Health Study Assets Research Program*. Washington (DC).
24. *Millennium Cohort Study*. (2020). San Diego, CA Retrieved from https://www.millenniumcohort.org/about.
25. DHA. (2019). *Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE) Fact Sheet*. Washington, DC Retrieved from https://health.mil/Reference-Center/Fact-Sheets/2019/08/21/ESSENCE.
26. DHA. (2020). *Aeromedical Services Information Management System (ASIMS) Fact Sheet*. Washington, DC Retrieved from https://health.mil/Reference-Center/Fact-Sheets/2020/09/02/ASIMS.
27. CDC. (2019). *Fourth Report on Human Exposure to Environmental Chemicals, Updated Tables*. Atlanta, GA: Department of Health and Human Services, Centers for Disease Control and Prevention Retrieved from https://www.cdc.gov/exposurereport/.
28. Lionetto, M., Caricato, R., & Giordano, M. (2019). Pollution Biomarkers in Environmental and Human Biomonitoring. *The Open Biomarkers Journal, 9*, 1-9.
29. Baumann, R., Gube, M., Markert, A., Davatgarbenam, S., Kossack, V., Gerhards, B., et al. (2018). Systemic serum amyloid A as a biomarker for exposure to zinc and/or copper-containing metal fumes. *J Expo Sci Environ Epidemiol, 28*(1), 84-91. doi:10.1038/jes.2016.86

30. Kapka-Skrzypczak, L., Cyranka, M., Skrzypczak, M., & Kruszewski, M. (2011). Biomonitoring and biomarkers of organophosphate pesticides exposure - state of the art. *Ann Agric Environ Med, 18*(2), 294-303.
31. Rothlein, J., Rohlman, D., Lasarev, M., Phillips, J., Muniz, J., & McCauley, L. (2006). Organophosphate pesticide exposure and neurobehavioral performance in agricultural and non-agricultural Hispanic workers. *Environ Health Perspect, 114*(5), 691-696. doi:10.1289/ehp.8182
32. Strelitz, J., Engel, L. S., & Keifer, M. C. (2014). Blood acetylcholinesterase and butyrylcholinesterase as biomarkers of cholinesterase depression among pesticide handlers. *Occup Environ Med, 71*(12), 842-847. doi:10.1136/oemed-2014-102315
33. Baxter, C. S., Hoffman, J. D., Knipp, M. J., Reponen, T., & Haynes, E. N. (2014). Exposure of firefighters to particulates and polycyclic aromatic hydrocarbons. *J Occup Environ Hyg, 11*(7), D85-91. doi:10.1080/15459624.2014.890286
34. Fent, K. W., Eisenberg, J., Snawder, J., Sammons, D., Pleil, J. D., Stiegel, M. A., et al. (2014). Systemic exposure to PAHs and benzene in firefighters suppressing controlled structure fires. *Ann Occup Hyg, 58*(7), 830-845. doi:10.1093/annhyg/meu036

## APPENDIX

**Table A1. Subjects Related to 1,690 Publications Not Included in Literature Analysis**

| Tertiary Search Term | Publications |
|---|---|
| hepatitis | 126 |
| sepsis/shock | 109 |
| diabetes | 103 |
| liver | 88 |
| wound | 82 |
| syndromes | 81 |
| kidney | 80 |
| pregnancy | 68 |
| TNF-a | 58 |
| allergy | 51 |
| tumor | 40 |
| diet | 27 |
| eye | 24 |
| Alzheimer's | 23 |
| heat/heat stress | 17 |
| schizophrenia | 17 |
| Parkinson's | 9 |
| COVID-19 | 6 |

PubMed query biomarker AND military AND tertiary search term.

**Table A2. Updates and Linked Sources of Open Access Databases**

| Database | Update Frequency (accessed 15 SEPT 2020) | Updated by | Linked/Source Databases |
|---|---|---|---|
| HMDB | Approx. every 3 years. V1 2006, V2 2009, V3 2012, V4 2018 | Wisehart Research Group | KEGG, PubChem, MetaCyc, ChEBI, PDB, UniProt, and GenBank |
| DrugBank | Daily. Downloads released quarterly. Latest version: 5.1.7 updated 07-02-2020 | Wisehart Research Group | Under Help tab, list of other databases |
| PubChem | nightly | Investigators submissions that undergo a review process | Over 500 data sources |
| ChEBI | monthly | ELIXIR Core Data Resource | IntEnz KEGG COMPOUND, PDBeChem, ChEMBL |
| OMIM | monthly | Investigators submissions that undergo a review process | External links tab under "help" links to 57 sources/databases |
| Exposome-explorer | Every few years. V1.0 was released in 2017 and V2.0 in 2019. | International Agency for Research on Cancer | Under the "about" tab there is a link to "other databases" |
| CDC National Biomonitoring Program | Infrequent. Last update for chemical factsheets is 07 APR 2017 | CDC | CDC, NIOSH, EPA |
| INCHEM | Undefined – likely based on source databases | IFCS, CCOHS | 13 data sources listed under "About INCHEM" link |
| MEB-Labs database | At least yearly, copyright on site is 2020 | Hosted and maintained by the University of Cincinnati | CDC, National Exposure Research Laboratory, etc. |
| CTD | At least once a month, last update 02 SEPT 2020 | CTD staff at NC State University | BioGRID, ChemIDplus®, cell ontology, gene ontology, KEGG, MeSH®, NCBI Gene, NCBI Taxonomy, PubMed®, Reactome, Uberon |

KEGG – Kyoto Encyclopedia of Genes and Genomes, ChEBI – Chemical Entities of Biological Interest, PDBe – Protein Data Bank in Europe, CDC – Centers for Disease Control and Prevention, NIOSH – National Institute for Occupational Safety and Health, EPA – Environmental Protection Agency, IFCS – Intergovernmental Program on Chemical Safety, CCOHS – Canadian Center for Occupational Health and Safety, BioGRID – Biological General Repository for Interaction Datasets, MeSH – Medical Subject Headings, NCBI – National Center for Biotechnology Information.

**ABBREVIATIONS**

| | |
|---|---|
| µM | micromolar |
| 711th HPW | 711th Human Performance Wing |
| ACGIH | American Conference for Governmental Industrial Hygienists |
| ACTUR | Automated Central Tumor Registry |
| AFHS | Air Force Health Study |
| AFI | Air Force Instruction |
| ASIMS | Aeromedical Services Information Management System |
| BEI | Biological Exposure Index |
| BioGRID | Biological General Repository for Interaction Datasets |
| CAC | Common Access Card |
| CAS | Chemical Abstracts Service |
| CCOHS | Canadian Center for Occupational Health and Safety |
| CCR | Consolidated Cancer Registry System |
| CDC | Centers for Disease Control and Prevention |
| ChEBI | Chemical Entities of Biological Interest |
| CTD | Comparative Toxicogenomics Database |
| DEERS | Defense Enrollment Eligibility System |
| DHA | Defense Health Agency |
| DHSS | Defense Health Service System |
| DoD | Department of Defense |
| DOEHRS | Defense Occupational and Environmental Health Readiness System |
| EPA | Environmental Protection Agency |
| ESSENCE | Electronic Surveillance System for Early Notification of Community-based Epidemics |
| HMDB | Human Metabolome Database |
| ICD | International Classification of Disease |
| IFCS | Intergovernmental Program on Chemical Safety |
| InCHEM | International Peer Reviewed Chemical Safety Information |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| MCS | Millennium Cohort Study |
| MEB-Labs | Measuring Environmental Biomarkers-Laboratories |
| MeSH | Medical Subject Headings |

| | |
|---|---|
| MHS | Military Health System |
| MHSRS | Military Health System Research Symposium |
| NCBI | National Center for Biotechnology Information |
| NIH | National Institute of Health |
| NIOSH | National Institute for Occupational Safety and Health |
| OMIM | Online Mendelian Inheritance in Man |
| OSHA | Occupational Safety and Health Administration |
| PDBeChem | Protein Data Bank in Europe |
| PM10 | Particulate matter with a diameter less than 10 micrometers |
| PM2.5 | Particulate matter with a diameter less than 2.5 micrometers |
| PON1 | Paraoxonasel1 gene mutation |
| PTSD | Post-traumatic stress disorder |
| Redox | Oxidation-reduction reaction |
| SNP | Single nucleotide polymorphism |
| TLV | Threshold Limit Value |
| WHO | World Health Organization |