Black Lung Incidence Study
Final Report

Date: **October 2023**

Prepared for: **U.S. Department of Labor, Chief Evaluation Office**

Contract Number: **1605DC-18-R-00044**

Prepared by: **Summit Consulting, LLC**
777 6th Street NW, Suite 520 | Washington, DC 20001
DUNS: 134631675 | Taxpayer ID: 95-4816438
CAGE Code: 3GZS9 | Unique ID: FNYSF7X4ELS5
www.summitllc.us

Disclaimer

This report was prepared for the U.S. Department of Labor (DOL), Chief Evaluation Office (CEO) by Summit Consulting, LLC, under contract number 1605DC-18-R-00044. The views expressed are those of the authors and should not be attributed to DOL, nor does mention of trade names, commercial products, or organizations imply endorsement of same by the U.S. Government.

Table of contents

EXECUTIVE SUMMARY **1**

1. INTRODUCTION **3**

2. DATA SOURCES AND METHODS **3**

2.1. Literature review **3**

2.2. Quantitative analysis **5**

2.2.1. Dataset scan *5*

2.2.2. Data cleaning and key definitions *6*

3. FINDINGS **12**

3.1. Black lung disease background **12**

3.1.1. Diseases associated with exposure to coal dust *12*

3.1.2. Non-occupational exposure to coal dust *15*

3.1.3. Diagnosing black lung disease and measuring incidence *15*

3.1.4. Black lung definition used for this study *18*

3.2. Black lung disease prevalence **18**

3.2.1. Black lung cases *18*

3.2.2. Black lung deaths *23*

3.3. Black lung prevalence across county types and populations of interest **28**

3.3.1. Context on Appalachian populations *29*

3.3.2. Context on the Navajo Nation *30*

3.3.3. Comparing prevalence across county types and populations of interest *31*

3.4. Limitations **36**

4. CONCLUSIONS AND NEXT STEPS **38**

4.1. Key takeaways by research question **38**

4.2. Potential next steps **39**

APPENDIX A SOURCES CITED **A-1**

APPENDIX B QUANTITATIVE ANALYSISR EFERENCES **B-1**

APPENDIX C MODELING BLACK LUNG PREVALENCE **C-1**

C.1: Analytic Methods **C-1**

C.2: Approach limitations **C-2**

C.3: Regression-estimated results **C-2**

List of figures

Figure 1: Coal county classification 8
Figure 2: U.S. Coal counties including borders for Appalachia and the Navajo Nation 11
Figure 3: Categorizing the health impacts of coal dust exposure, from narrowest to broadest interpretation........ 12
Figure 4: Nationwide prevalence of coal workers' pneumoconiosis among coal miners with 25+ years of experience, from literature review findings 14
Figure 5: Cumulative black lung cases versus number of coal mines (1983, 2020) 20
Figure 6: Cumulative black lung cases versus amount of coal production in short tons (1983, 2020) 20
Figure 7: Cumulative black lung cases versus number of coal employees (1986, 2020) 21
Figure 8: Cumulative black lung cases versus weekly hours of coal dust exposure (hours, 1970–2020) .. 21
Figure 9: Cumulative black lung cases versus number of households burning coal (2021) 22
Figure 10: Cumulative black lung deaths versus number of coal mines (1983, 2020) 24
Figure 11: Cumulative black lung deaths versus amount of coal production in short tons (1983, 2020) .. 25
Figure 12: Cumulative black lung deaths versus number of coal employees (1986, 2020) 25
Figure 13: Cumulative black lung deaths versus weekly hours of coal dust exposure (hours, 1970–2020) .. 26
Figure 14: Cumulative black lung deaths versus number of households burning coal (2021) 26
Figure 15: Cumulative black lung deaths (1999–2020) 27
Figure 16: Cumulative black lung death rates (per 1,000 residents in 2020) (1999–2020)........ 27
Figure 17: Cumulative black lung death rates among all ICD-10 diagnoses (1999–2020)........ 28
Figure 18: Cumulative black lung death rates attributable to CWP (1999–2020) 28

List of tables

Table 1: Summary of key findings 1
Table 2: List of search terms 4
Table 3: Inclusion and exclusion criteria........ 4
Table 4: Data sources retained for black lung quantitative analyses........ 5
Table 5: U.S. counties, coal versus non-coal........ 9
Table 6: U.S. counties, scale of coal county status........ 9
Table 7: Coal counties, current versus former 10
Table 8: Black lung disease diagnostic codes........ 16
Table 9: Cumulative black lung cases by severity across the entire collection period, 1970-2014 19
Table 10: Top ten counties in terms of cumulative black lung cases and rates, 1970-2014 22
Table 11: Summary statistics of cumulative black lung deaths by diagnosis code across the entire collection period, 1999-2020 24
Table 12: Comparison of difference in black lung counts and rates, coal versus non-coal counties........ 32
Table 13: Comparison of difference in black lung counts and rates, current versus former coal counties34

Table 14: Comparison of difference in black lung counts and rates, Appalachia versus other counties ... 34
Table 15: Comparison of difference in black lung counts and rates, Navajo Nation versus other counties 35
Table 16: Comparison of difference in black lung counts and rates, surface versus underground mines. 35
Table 17: Comparison of difference in black lung counts and rates, residential coal use 36
Table 18: Key takeaways by research question 38

Executive summary

Background. A key part of the mission of the Mine Safety and Health Administration (MSHA) is to protect coal miners from death, illness, and injury. To support this mission, Summit Consulting (Summit) conducted this Black Lung Incidence Study to address the following research questions (RQs):

- **RQ1:** What is the rate of black lung disease across the United States?
  - RQ1.1: What is the total number and rate per 1,000 residents of black lung cases?
  - RQ1.2: What is the total number and rate per 1,000 residents of black lung deaths?
- **RQ2:** How does black lung incidence compare between current, former, and non-coal mining communities?
- **RQ3:** Are black lung cases and deaths more prevalent in the Navajo Nation or Appalachia than other parts of the United States?
- **RQ4:** How does residential coal burning correlate with black lung cases and deaths?

Data sources and methods. To answer the study research questions, Summit conducted the following activities:

- A literature review to explore the current state of knowledge on black lung definitions, diagnosis, and measurement; coal mining and black lung disease; the link between residential coal use and black lung disease; black lung incidence in Appalachia; and black lung incidence in Navajo Nation populations.
- A dataset scan that identified publicly available data for analysis from the following sources:
  - U.S. Census Bureau American Community Survey;
  - U.S. Energy Information Administration;
  - U.S. Census Bureau County Business Patterns;
  - Center for Disease Control and Prevention (CDC) Wide-Ranging ONline Data for Epidemiologic Research;
  - CDC Enhanced Coal Workers' Health Surveillance Program; and
  - MSHA Mines Data Set.
- A statistical analysis (descriptive and inferential) of publicly available data to identify black lung incidence, determine whether the causes for higher rates of black lung disease can be disaggregated, and design a series of predictive models to predict the number of black lung cases or deaths across the United States.

Key findings. The key findings from the Black Lung Incidence Study presented in this final report are summarized in Table 1 below.

Table 1: Summary of key findings

RQ1: What is the rate of black lung disease across the United States?
• On average, there are 4.34 cumulative cases (1970-2014) and 3.44 cumulative deaths (1999-2020) attributable to black lung disease per county across the United States. These results are counts and therefore not subject to population size.
• The prevalence of black lung disease is highly concentrated in specific areas of the country, such as Appalachia, where the statistics rise to 28.79 cumulative cases and 10.88 cumulative deaths, on average. 103 counties out of 3,136 report black lung cases during the entire collection period (1970-2014) while 333 counties report black lung deaths (1999-2020). These results are counts and not subject to population size.
• Coal workers' pneumoconiosis (CWP) prevalence has been increasing in the United States since the 1990s.

• Undercounts of black lung disease in the United States may be attributed to factors including low uptake of Coal Workers' Health Surveillance Program screenings and low compensation approval rates.
RQ1.1: What is the total number and rate per 1,000 residents of black lung cases?
• There are 0.11 black lung cases per 1,000 U.S. residents, or 11 cases per 100,000.*
RQ1.2: What is the total number and rate per 1,000 residents of black lung deaths?
• There are 0.04 black lung deaths per 1,000 U.S. residents, or 4 cases per 100,000.*
RQ2: How does black lung incidence compare between current, former, and non-coal mining communities?
• Black lung disease is most prevalent among coal counties,^ especially those that have maintained coal-mining practices since the 1970s and 1980s.
• Former coal counties have significantly lower black lung rates than current coal counties.
• Non-coal counties rarely report any cases or deaths attributable to black lung disease. Deaths in these counties most often come from asbestosis, which may risk overcounting.
RQ3: Are black lung cases and deaths more prevalent in the Navajo Nation or Appalachia than other parts of the United States?
• Appalachia has significantly more black lung cases and deaths than the rest of the United States, both overall and accounting for population size.
• Zero cases and just 10 deaths were reported in the Navajo Nation during the respective collection periods (1970-2014 and 1999-2020). Therefore, hypothesis testing between the Navajo Nation and other regions proved inconclusive. These results may be a function of underreporting in the Navajo Nation.
RQ4: How does residential coal burning correlate with black lung cases and deaths?
• There is a correlation between residential coal use and black lung cases and deaths. However, this correlation does not imply a causal relationship between residential use and black lung disease.
• There are significantly more cases and deaths from black lung in residential coal burning areas than other parts of the United States.
• Residential coal burning is likely a confounding factor with other black lung characteristics, such as the number of local underground mines or family members working in the coal industry.

* Results are cumulative across the respective data collection period (i.e., black lung cases 1970-2014, black lung deaths 1999-2020). Population estimates are based on 2014 and 2020 data, respectively.

^ Summit classified a county as a coal county if any of the following four criteria are met: county has 1+ mines in either 1983 or 2020, county produced 1+ short tons of coal in 1983 or 2020, there is 1+ coal miners residing in the county in 1986 or 2020, or workers in the county are exposed to an average of 1+ hours to coal mine dust each week between 1970 and 2020.

Note: Key findings were derived based on literature review sources (see Appendix A) as well as publicly available data sources for quantitative analyses (see Section 2.2.1).

Conclusions and potential next steps. The results of this study show that black lung disease due to unsafe practices in coal mining, residential coal burning, and air pollution through coal processing and transportation remains at concerning levels in the United States. In addition, Appalachia has significantly more black lung cases and deaths than the rest of the United States, both overall and controlling for population size. This study identified areas of potential future research to help MSHA understand the impact of coal on health:

- Future research to address the lack of public health data available for analysis in the Navajo Nation;
- A follow-on study with similar research questions that focus specifically on silicosis (a lung disease associated with exposure to silica dust)—rather than on black lung disease more broadly—in response to the recently announced proposed rule change (related to 30 CFR § 56, 57, 70, 71, 72, and 90) to address health hazards from silica dust exposure (MSHA 2023). Because this rule change was announced after the literature review and statistical analysis for this report had been completed, this report includes silicosis in combination with other respiratory diseases but does not examine silicosis individually.

1. Introduction

Study scope and research questions. The Federal Mine Safety and Health Act of 1977 chartered the Mine Safety and Health Administration (MSHA), whose mission is to "prevent death, illness, and injury from mining and promote safe and healthful workplaces for U.S. miners" (29 U.S.C. § 557a; 30 CFR § 72.1; 30 CFR § 72.510). To support this mission, Summit Consulting (Summit) conducted this Black Lung Incidence Study to examine black lung incidence in the United States, exploring both cases and deaths. Within this scope, the study examined whether black lung incidence is higher among specific subpopulations of interest, including miners, mining communities, the Navajo Nation, and residents of Appalachia.[1]

> ***Black lung disease*** is an umbrella term that refers to a variety of respiratory diseases such as coal workers' pneumoconiosis and silicosis. These diseases are caused by exposure to coal dust and have no cure.

This study addressed four research questions (RQs):

- **RQ1:** What is the rate of black lung disease across the United States?
  - **RQ1.1:** What is the total number and rate per 1,000 residents of black lung cases?
  - **RQ1.2:** What is the total number and rate per 1,000 residents of black lung deaths?
- **RQ2:** How does black lung incidence compare between current, former, and non-coal mining communities?
- **RQ3:** Are black lung cases and deaths more prevalent in the Navajo Nation or Appalachia than other parts of the United States?
- **RQ4:** How does residential coal burning correlate with black lung cases and deaths?

Summit conducted a literature review and quantitative analysis using publicly available data to address these research questions (see Table 4 in Section 2.2.1 for a description of the publicly available datasets used). The literature review focused on black lung disease diagnosis and measurement, coal mining, residential coal burning, and black lung incidence in Appalachian populations and the Navajo Nation. Summit also conducted statistical analysis (descriptive and inferential) to identify black lung incidence prior to death and determine whether the causes for higher rates of black lung disease can be disaggregated, as detailed in Section 3.

2. Data sources and methods

2.1. Literature review

Summit conducted a literature review to explore the current state of knowledge on black lung definitions, diagnosis, and measurement; coal mining and black lung disease; the link between residential coal use and black lung disease; black lung incidence in Appalachia; and black lung incidence in Navajo Nation populations.

[1] Appalachia is a geographical region centered around the Appalachian Mountains covering West Virginia and parts of 12 other states (West Virginia University 2022). The Navajo Nation is located in northeastern Arizona, northwestern New Mexico, and southeastern Utah (Navajo Nation, accessed September 29, 2023.).

The team followed a rigorous methodology and documentation process that included (1) developing key search terms, (2) identifying inclusion and exclusion criteria, (3) conducting the initial literature search, (4) performing analysis, and (5) completing additional literature searches as needed.

The key search terms (with variations in parentheses) are shown in Table 2. Searches were conducted via Google Scholar using combinations of these terms. For example, the team searched for "Navajo Nation" and "home coal burning," rather than simply searching "Navajo Nation."

Table 2: List of search terms

- **Appalachia** (Appalachian)
- **Coal mining communities**
- **Household cooking** (searched in combination with "coal")
- **Black lung** (pneumoconiosis, coal workers' pneumoconiosis, emphysema, silicosis, bronchitis, respiratory illness, respiratory disease, lung cancer)
- **Coal workers' pneumoconiosis definition** (coal workers' pneumoconiosis medical codes)
- **Indoor air pollution**
- **Black lung definition** (black lung medical diagnosis codes)
- **Home coal burning** (residential coal burning, residential coal combustion)
- **Navajo Nation**

In addition to searching the terms in the table above, Summit also solicited relevant sources from the project subject-matter expert, Dr. Robert Finkelman, a research professor of geosciences at the University of Texas at Dallas with expertise on the social and health impacts of coal and black lung disease. Dr. Finkelman's research has focused on the health impacts of coal, including residential coal use.

Summit then developed inclusion and exclusion criteria to further focus the search, shown in Table 3. The inclusion criteria are requirements that must be met for a source to be considered for analysis. The exclusion criteria are characteristics that lead to the omission of a source for analysis.

Table 3: Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
• Source is relevant to one of the following: ○ Black lung disease resulting from residential coal burning ○ Black lung disease resulting from a coal-related occupation ○ Contributes to black lung disease definition or medical code discussion	• Source is older than 20 years (2002 or earlier); this exclusion criterion was applied with flexibility and did not apply to federal legislation.

Following the search, 124 sources were saved for inclusion in the review. Summit reviewed all sources for relevant information and synthesized findings to focus on topics of interest. The list of sources cited in this report is provided in Appendix A.

2.2. Quantitative analysis

2.2.1. Dataset scan

Summit conducted a dataset scan to understand and identify what data are publicly available to support research questions. The datasets queried during this process were aggregated into a single file that was used for the quantitative analyses.

As the team explored publicly available databases, the dataset scan served three primary purposes:

1. Identify sources that track black lung cases or deaths across time;
2. Determine how these sources define black lung disease using International Classification of Diseases (ICD)-10 diagnosis codes; and
3. Determine coal-county status using county-level regional classifications and descriptive characteristics of coal use and production.

Summit reviewed a series of federal databases to access aggregated public health data across time in the United States. Specifically, Summit reviewed statistics from the U.S. Centers for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) database to determine county-level death data from black lung diseases (CDC 2023).[2] In addition, the research team reviewed and queried data from the Enhanced Coal Workers' Health Surveillance Program (ECWHSP) database, another CDC-backed worker monitoring program that tracks living black lung case data via X-rays (CDC NIOSH 2020).

Table 4 below summarizes the databases from which Summit gathered data for analysis.[3] Because the structure, key definitions, and regularity of collection across datasets is not uniform, Summit made key judgments in order to construct a single dataset that best answered all the research questions within the time and budget constraints. Section 2.2.2 discusses the data cleaning process (i.e., data quality assurance) and this decision-making process.

Table 4: Data sources retained for black lung quantitative analyses

Title	Description	Years Included	Relevant Data
CDC Wide-ranging ONline Data for Epidemiologic Research (WONDER)	Aggregated health statistics including AIDS, cancer, STDs, and morbidity and mortality	1999*–2020^	Total black lung deaths by ICD-9/ICD-10 diagnosis
CDC Enhanced Coal Workers' Health Surveillance Program (ECWHSP)	Coal mining medical monitoring database sponsored by the CDC National Institute for Occupational Safety and Health	1970*–2014^	Total reported black lung cases, black lung severity, mine type

[2] Black lung diseases are defined according to a series of diagnosis codes discussed in Section 3.1.3.
[3] Links to all quantitative data sources are provided in Appendix B.

Title	Description	Years Included	Relevant Data
U.S. Census American Community Survey, 1-year estimates	Annual demographics survey that includes information such as population, number of total households, and proportion of homes using coal for heating or cooking	2014; 2020; 2021^	County name, state, FIPS code, population, total deaths, total households, households using coal or coke as fuel
U.S. Census County Business Patterns (CBP)	Annual report providing labor statistics by industry (North American Industry Classification System (NAICS) or Standard Industrial Classification (SIC) codes[4])	1986*; 2020^	Total mining employees
U.S. Energy Information Administration (EIA)	Annual statistical report that analyzes and projects energy production and usage including coal	1983*; 2020	Coal mining regions, coal mine production
Mine Safety and Health Administration (MSHA) Mines Data Set	Panel dataset of all coal mines under MSHA jurisdiction as well as active status	1970*–2020	Number of coal mines, mine type (underground or surface), coal type, production hours, mine activity status, MSHA district

* Indicates the earliest available data.
^ Indicates the most currently available data at the time of collection.

2.2.2. Data cleaning and key definitions

After collecting the publicly available data for analysis, Summit merged the data into a single analytic file. The first step of this process was to organize and classify black lung data points according to how they are collected. CDC WONDER deaths data are county-level data describing the cumulative number of deaths between 1999 and 2020 attributable to a specific ICD-9 or ICD-10[5] diagnosis code. Summit segmented these data from CDC WONDER according to black lung diagnoses identified using available literature and consultation with black lung experts, resulting in a county-level dataset of cumulative deaths across the seven diagnoses found in Table 8 located in Section 3.1.3.

Black lung cases, on the other hand, are only tracked by ECWHSP according to a single diagnosis code: ICD-9 code 500 (ICD-10 code J60: coal workers' pneumoconiosis). The collection process for these data is not only a more restrictive definition of black lung disease than Table 8 below, but it also covers a slightly different time period (1970–2014). Like the black lung deaths data, these incidence data are also cumulative across time. However, the black lung case data from ECWHSP classify each case according to its severity, ranging from 1 (least severe) to 3 followed by progressive massive fibrosis (PMF, the most

[4] The research team reviewed NAICS codes and SIC codes to classify coal-mining business patterns. Coal mining NAICS codes are 2121, 212111, 212112, and 212113. SIC codes are 1211, 1213, 122, 1221, 1222, 123, 1231, 124, and 1241.
[5] Note: ICD-9 diagnosis code schemes were replaced beginning October 1, 2015. ICD-10 diagnoses have been the default coding schema moving forward and are still current at the time of this publication.

severe), according to the International Labour Organization.[6] In order to gather as much usable information as possible for our research, the team decided to keep these data separate from black lung deaths so we could better interpret the results and understand the differences between these two metrics.

The researchers brought in population statistics from the 2014, 2020, and 2021 iterations of the U.S. Census Bureau's American Community Survey (United States Census Bureau 2023).[7] The team also leveraged household-level information including the number of households burning coal residentially from the 2021 iteration of the survey. These data also allowed the team to understand the number and share of households in each county that use coal (or coke, a closely related fossil fuel) in the home for residential heating or cooking.

Summit also gathered the most recently available Census data on county business patterns[8] (2020) to understand the coal industry's influence in each county (United States Census Bureau 2022b). Census only began collecting this data in 1986, so this earliest iteration was used as a historical comparator for all black lung statistics from 1970 onward. Finally, the team included geographic identifiers such as "Appalachia" for coal-producing regions from the U.S. Energy Information Administration (EIA)[9] as well as mine-level information recorded quarterly from MSHA[10] from 1970-2021 (United States Energy Information Administration 2023, United States Department of Labor 2023).

The resulting aggregated dataset was composed entirely of publicly available data collected from various federal programs. Summit used this dataset to define coal-county status and populations of interest to support the research questions. Specifically, the team composed county-level definitions for those associated with the coal industry ("coal counties"), the Appalachian and Navajo Nation territories, and those whose residents burn coal residentially.

> The team defines "coal counties" as counties impacted by the coal industry (i.e., existence of coal mines, coal mine employees, coal production) based on publicly available data.

Defining coal counties. In this report, "coal county" is a binary indicator to define counties impacted by the coal industry. Specifically, this binary outcome is determined according to the following four criteria: coal mines, coal production, coal miners, and exposure to coal dust. Under this methodology, a county was classified as a coal county when *any* of the following four criteria are met:

---

[6] International Labour Organization. 2002. "Guidelines for the Use of the ILO International Classification of Radiographs of Pneumoconioses." https://www.ilo.org/wcmsp5/groups/public/@ed_protect/@protrav/@safework/documents/genericdocument/wcms_861207.pdf

[7] The American Community Survey provides information on the social and economic needs across communities each year by sampling approximately 3.5 million addresses.

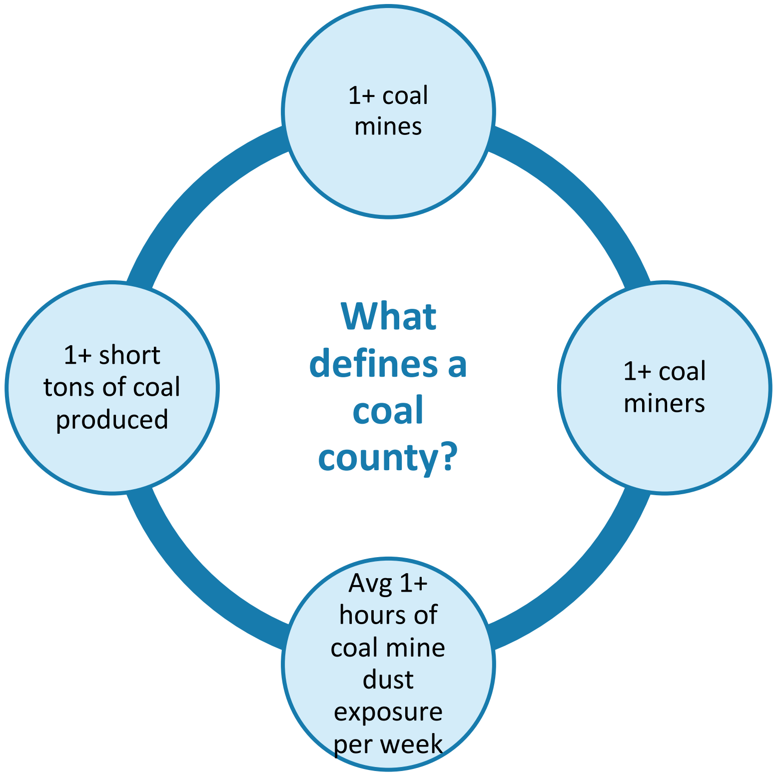
[8] The County Business Patterns (CBP) is an annual series that provides economic data according to industry (NAICS and SIC codes). Among other things, CBP includes information about the number of firms, workers, and associated payroll.

[9] The EIA annual coal report provides data on domestic coal production as well as the number of mines and employment. This report leverages data from MSHA's mine dataset, which was also used for this study.

[10] MSHA's Mine Data Set provides quarterly information on all U.S. mines under MSHA jurisdiction. It includes relevant information on active status, production levels, and type(s) of coal produced. Each mine has a unique ID to track statistics across time. This report aggregates all mines in a county from individual mine-level results.

1. The county has 1+ mines in either 1983 or 2020;[11]
2. The county produced 1+ short tons of coal in 1983 or 2020;
3. There is 1+ coal miners residing in the county in 1986 or 2020; or
4. Workers in the county are exposed to an average of 1+ hours to coal mine dust each week between 1970 and 2020.

Figure 1: Coal county classification



This definition was designed to capture as many relevant counties as possible without being overly stringent. However, in most cases, if one coal county criterion was met, other criteria were also met (e.g., counties with a coal mine likely also produce coal and/or have employees who are exposed to coal dust). Of the 3,136 U.S. counties studied[12], 286 (9%) are classified as coal counties according to one or more of these parameters.

[11] EIA and Census Bureau County Business Patterns datasets are collected at specific points in time. Given the differences in collection intervals and data availability, the team opted to select the oldest and most current points in time as references. Constructing a panel dataset across the seven unique data sources would have exceeded the project's budget and time constraints.

[12] Note: As of 2020, there are 3,143 counties according to the U.S. Census Bureau. However, 7 counties either had completely missing data or merged with other counties to generate statistics. Merged counties occur in Alaska only.

Table 5: U.S. counties, coal versus non-coal

Name	Count	Percent
Coal	286	9%
Non-coal	2,850	91%
TOTAL	**3,136**	**100%**

Note: Coal county status is based on an aggregate definition according to each county's number of coal mines, amount of coal produced (in short tons), number of coal workers residing in the county, and the amount of coal dust exposure (in hours) for those workers. This underlying data stems from U.S. Energy Information Administration (EIA) annual coal report and U.S. Department of Labor (DOL) Mine Safety and Health Administration (MSHA) mines data set. All definitions and corresponding data results presented in the table are original and were created as part of this research.

In addition to a binary definition of coal counties, Summit also classified each county's coal level on a sliding scale according to percentiles within each criterion. This means that each criterion has its own independent scale on which each county is assessed. The scale includes the titles "no coal," "very low," "low," "medium," "high," and "very high." These individual ratings for each criterion were compared to determine a county's overall association with coal along a sliding scale. [13] In order to best identify counties closest associated with the coal industry, each county was given an overall classification according to the highest individual criterion score. For example, counties with any individual criterion in the "very high" category were given an overall classification of "very high."

Under these parameters, 137 counties (4%) are considered "very high" coal counties. This amount is roughly equal to the sum of all other types of coal counties (149), meaning that these 137 counties are the most closely associated with the U.S. coal industry. The breakdown of each scaled definition can be found in Table 6, below.

Table 6: U.S. counties, scale of coal county status

Coal Status	Count	Percent
Very high	137	4%
High	32	1%
Medium	36	1%
Low	39	1%
Very low	42	2%
Non-coal	2,850	91%
TOTAL	**3,136**	**100%**

Note: Coal county status is based on an aggregate definition according to each county's number of coal mines, amount of coal produced (in short tons), number of coal workers residing in the county, and the amount of coal dust exposure (in hours) for those workers. This underlying data stems from U.S. Energy Information Administration (EIA) annual coal report and U.S. Department of Labor (DOL) Mine Safety and Health Administration (MSHA) mines data set. All definitions and corresponding data points in the table are original and were created as part of this research.

Coal counties were further dissected according to their historical status, either "current" or "former." Current coal counties are classified according to their most recent coal status (e.g., 2020 coal production), while former coal counties rely on both the most recent statistics as well as historical figures. For example, Muskingum County, Ohio produced more than 3,200 short tons of coal in 1983. By 2020, Muskingum's coal production ceased entirely, defining it as a former coal county. Of the 286 coal counties, 172 are considered current and 114 are former, as shown in Table 7 below. Interestingly, of the 172 current coal counties, 154 also have some historical precedence with coal. This means that only

[13] Scalar labels are defined using percentiles for each unique parameter. The exact thresholds for each parameter can be found in Table B.2 located in Appendix B.

18 counties could be considered "new" coal counties under these parameters (i.e., the county's relationship with the coal industry began after 1986).

Table 7: Coal counties, current versus former

Historic Status	Count	Percent
Current coal	172	60%
Former coal	114	40%
TOTAL	**286**	**100%**

Note: Coal county status is based on an aggregate definition according to each county's number of coal mines, amount of coal produced (in short tons), number of coal workers residing in the county, and the amount of coal dust exposure (in hours) for those workers. This underlying data stems from U.S. Energy Information Administration (EIA) annual coal report and U.S. Department of Labor (DOL) Mine Safety and Health Administration (MSHA) mines data set. Current coal counties are based on the most currently available statistics (i.e., 2020 data) created for this research. Former coal counties depend on both earliest (1970, 1983, and 1986, depending on the metric) and most currently available statistics for classification. All definitions and corresponding data points in the table are original and were collected as part of this research.

Defining Appalachia and the Navajo Nation. EIA maintains a list of definitions for coal-producing regions across the United States, including Appalachia.[14] Summit adopted definitions from EIA's coal glossary and assigned each relevant county one of the following regional identifiers to define the following geographic groups in the analytic file[15]:

1. Appalachian Region: Aggregated region in the United States made up of:
    a. North Appalachia: Consists of Maryland, Ohio, Pennsylvania, and Northern West Virginia.
    b. Central Appalachia: Consists of Eastern Kentucky, Virginia, Southern West Virginia, and the Tennessee counties of: Anderson, Campbell, Claiborne, Cumberland, Fentress, Morgan, Overton, Pickett, Putnam, Roane, and Scott.
    c. South Appalachia: Consists of Alabama, and the Tennessee counties of: Bledsoe, Coffee, Franklin, Grundy, Hamilton, Marion, Rhea, Sequatchie, Van Buren, Warren, and White.
2. Interior Region (with Gulf Coast): Consists of Arkansas, Illinois, Indiana, Kansas, Louisiana, Mississippi, Missouri, Oklahoma, Texas, and Western Kentucky.
3. Illinois Basin Region: Consists of Illinois, Indiana, and Western Kentucky.
4. Western Region: Consists of Alaska, Arizona, Colorado, Montana, New Mexico, North Dakota, Utah, Washington, and Wyoming.
5. Powder River Basin Region: Consists of the Montana counties of Big Horn, Custer, Powder River, Rosebud, and Treasure and the Wyoming counties of Campbell, Converse, Crook, Johnson, Natrona, Niobrara, Sheridan, and Weston.

Given that the Navajo Nation is its own self-governing entity and does not directly align with U.S. county lines, Summit opted to define Navajo Nation counties according to the six U.S. counties that either partially or totally reside within the Navajo Nation's borders using U.S. Census data. These six counties include:

1. Apache County, AZ;
2. Coconino County, AZ;
3. Navajo County, AZ;

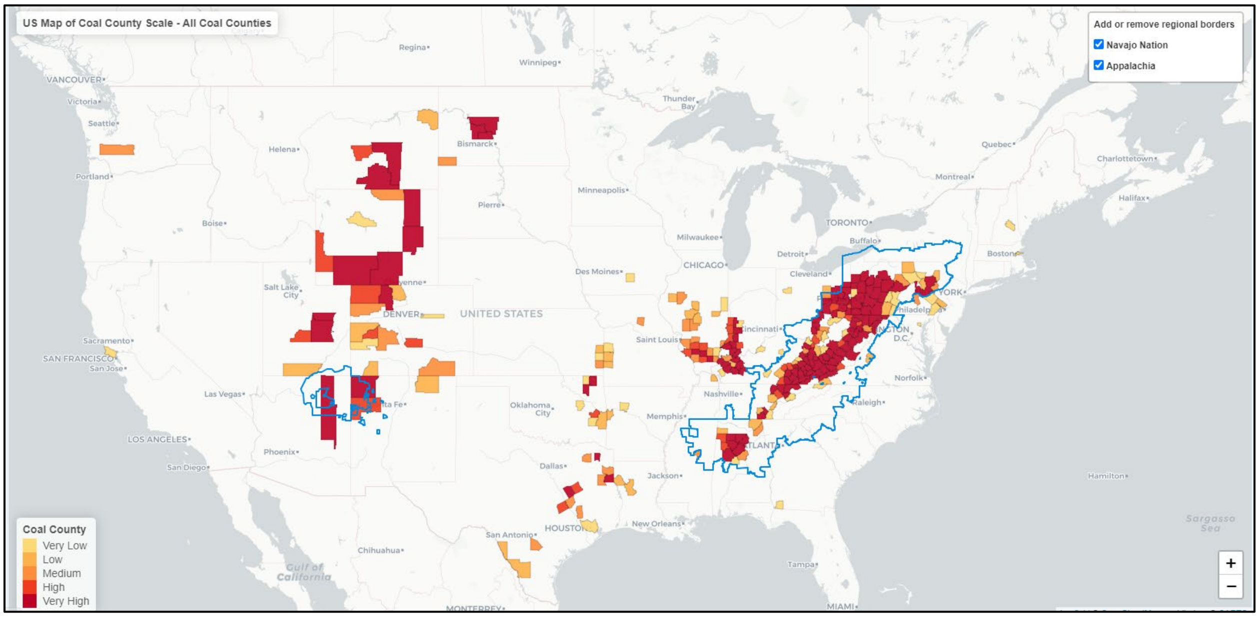
[14] U.S. Energy Information Administration. "Glossary." https://www.eia.gov/tools/glossary/?id=coal. Accessed August 3, 2023.
[15] Appendix B includes a breakdown of county counts and percentages into their respective geographic categories.

4. McKinley County, NM;
5. San Juan County, NM; and
6. San Juan County, UT.

Figure 2 below provides a visual representation of coal counties across the United States overlaid by border of the Navajo Nation and Appalachia regions. The Navajo Nation and Appalachia borders are in blue on the left (west) and right (east), respectively.

Figure 2: U.S. Coal counties including borders for Appalachia and the Navajo Nation

Map of the continental United States categorizing counties based on coal-county status. Coal county categories include "Very Low", "Low", "Medium", "High", and "Very High." Borders for the Navajo Nation and Appalachia are included in the map.



Note: Coal county status is based on an aggregate definition according to each county's number of coal mines, amount of coal produced (in short tons), number of coal workers residing in the county, and the amount of coal dust exposure (in hours) for those workers. This underlying data stems from U.S. Energy Information Administration (EIA) annual coal report and U.S. Department of Labor (DOL) Mine Safety and Health Administration (MSHA) mines data set. All definitions and corresponding data points in the table are original and were collected as part of this research.

Defining counties with residential coal use. Summit classified counties as "residential use" counties according to whether their residents burn coal as the primary source of fuel in the home. The U.S. Census Bureau's 2021 iteration of the American Community Survey asks respondents which fuel is most used in their homes for heat. Using this data, less than 0.1% of all households in the country report coal as the primary source of residential heat. Denali County, Alaska, the highest individual county in terms of residential coal use, reports roughly 17% of households burning coal as its primary fuel source.

As a means to gather as much usable data as possible, Summit created an inclusive definition for its quantitative analysis. Counties where one or more households report residential coal burning as its primary fuel source are assumed for all households in the county – that entire county is considered a residential coal-use county. In total, this inclusive definition captures roughly a third (1,105) of counties in the dataset.

3. Findings

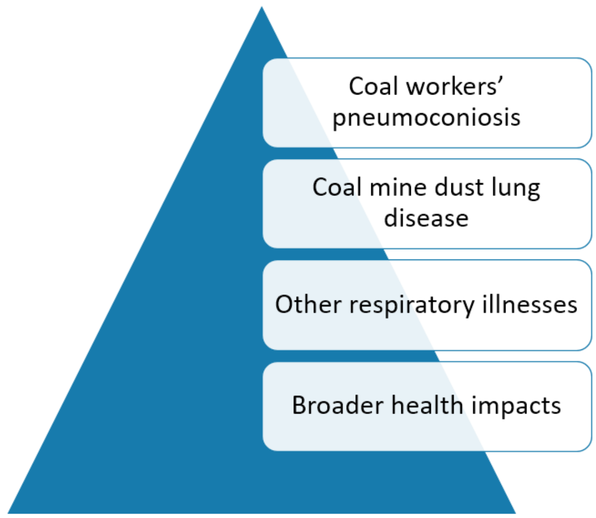
3.1. Black lung disease background

"Black lung disease" is an umbrella term that can refer to a variety of diseases associated with respirable coal mine dust; variations in which specific diseases are included in measurements of black lung case and death rates can create inconsistency across measures of nationwide prevalence. This section will provide an overview of the respiratory diseases associated with exposure to coal dust, incidence measurement challenges and potential undercounts, and the specific diseases (and associated medical diagnosis codes) included in our quantitative analysis.

3.1.1. Diseases associated with exposure to coal dust

There are many specific diseases that are associated with exposure to coal dust, whether due to coal mine employment or nonoccupational exposure. This range of diseases is organized in Figure 3 from the narrowest possible definition of black lung disease at the top of the pyramid to the broadest interpretation at the bottom.

Figure 3: Categorizing the health impacts of coal dust exposure, from narrowest to broadest interpretation



Coal worker's pneumoconiosis represents the narrowest interpretation of black lung disease. Coal mine dust lung disease captures a broader spectrum of disease, which can include CWP, silicosis, mixed-dust pneumoconiosis, PMF, dust-related diffuse fibrosis (DDF), COPD, and chronic airway diseases. Other respiratory diseases associated with coal dust exposure include lung cancer, asthma, decreased lung function, and acute lower respiratory infections. Finally, coal dust exposure is also associated with broader health impacts including low birthweight in newborn infants, increased infant mortality, neurological effects, mental illness, cataracts and immune system impairment.

Coal workers' pneumoconiosis (CWP). Historically, high rates of pneumoconiosis have occurred in mining settings due to chronic coal dust exposure (Perret et al. 2017; Patra et al. 2016; M. H. Ross and Murray 2004). The narrowest interpretation of black lung disease is CWP, a lung disease caused by chronic inhalation of coal dust for which there is no cure (Paul, Adeyemi, and Arif 2022; Cecil 2021; Finkelman, Wolfe, and Hendryx 2021; Arif et al. 2020; Royal 2019; Zosky et al. 2016; Laney and Weissman 2014; Lockwood 2012, 52, 125–6; Huang et al. 2006). CWP is shown at the top of Figure 3. Inhaled coal dust accumulates in the lungs, causing inflammation and the formation of lesions called coal macules (McCunney, Morfeld, and Payne 2009). Although the link between CWP and coal dust is well established and accepted by the medical community, no definitive scientific explanation for this link has been identified (Song et al. 2022; Sun, Kinsela, and Waite 2022; Zosky et al. 2021; Harrington et al. 2012; Cohn et al. 2006; Kuempel et al. 2003).[16]

CWP cases can be classified as simple or complicated based on the size of the lesions (Arif et al. 2020; NIOSH 2020). Simple CWP is characterized by small nodules (1 to 2 mm) made up of immune and

[16] Quartz, pyrite, and bioavailable iron have been theorized as potential causal factors of illness (Harrington et al., 2012; Schoonen et al. 2010; Cohn et al. 2006).

inflammatory cells, collagen fibers, and coal dust (Arnold 2016). Symptoms can include chronic cough, increased phlegm production, and shortness of breath (Arnold 2016). Simple CWP can also be asymptomatic, which could lead to undercounts of the true number of cases if asymptomatic cases are not diagnosed (Paul, Adeyemi, and Arif 2022; Zosky et al. 2016; Hendryx et al. 2013; Finkelman et al. 2002). Simple CWP can develop into progressive massive fibrosis (PMF)—otherwise known as complicated CWP—as the size of coal nodules increase and begin to tear the surrounding lung tissue (Zosky et al. 2016; McCunney, Morfeld, and Payne 2009; Finkelman et al. 2002). As the most severe form of CWP, PMF is a “rapidly progressive and often fatal disease” (Cecil 2021; Reynolds et al. 2018; Lockwood 2012, 126). The risk of developing PMF increases the longer a coal miner has worked, due to a direct relationship between the amount of coal dust inhaled and the incidence and severity of CWP (Lockwood 2012, 52; Finkelman et al. 2002).

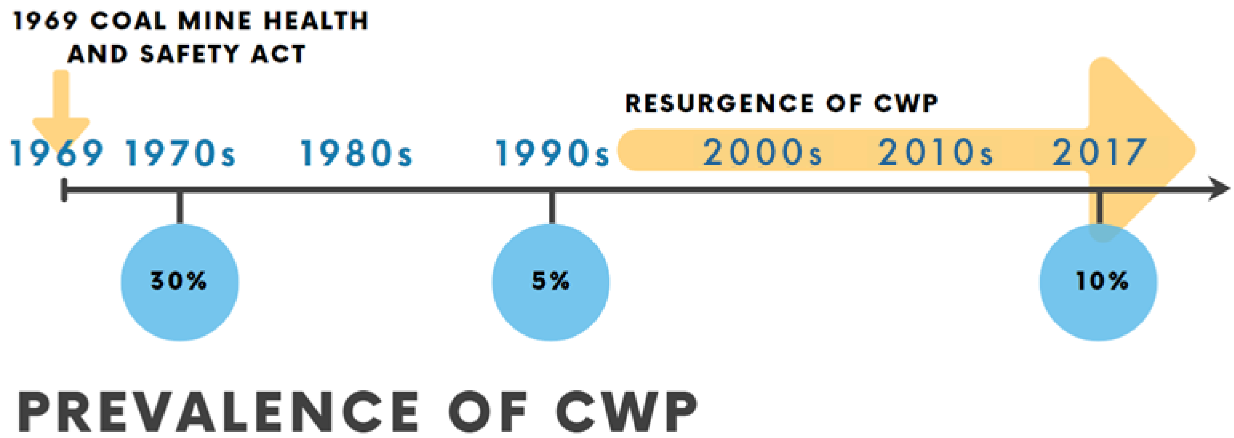
Coal miners’ exposure to dust can depend on several factors such as mining methods and job duties (NIOSH 2020). For example, underground miners generally have greater exposure than surface miners (NIOSH 2020; Huang et al. 2006). Additionally, duties such as coal transfer points on conveyor systems may lead to higher dust exposure (Huang et al. 2006). Under Part 90 of the Federal Coal Mine Health and Safety Act, coal miners who have been diagnosed with pneumoconiosis are eligible for a transfer to a different mine or position without discrimination, including pay reduction or termination (30 CFR § 90.103).

Understanding the (Enhanced) Coal Workers’ Health Surveillance Program. The Coal Workers’ Health Surveillance Program (CWHSP) conducts mobile black lung screenings across the United States through chest X-rays and spirometry tests with miners. The Enhanced Coal Workers’ Health Surveillance Program (ECWHSP) collates the X-rays from the CWHSP and also draws data from the Coal Workers’ X-ray Surveillance Program.

Figure 4 shows literature review findings on the nationwide prevalence of CWP among coal miners over time, starting shortly after the 1969 Coal Mine Health and Safety Act.[17] The prevalence of the disease decreased after this legislation’s provisions reducing dust exposure limits (Shriver and Bodenhamer 2018, Laney and Weissman 2014). However, there has been a rise in prevalence since the 1990s (Laney and Attfield 2014; Laney and Attfield 2010). A study conducted by the surveillance branch of the National Institute for Occupational Safety and Health (NIOSH) Division of Respiratory Disease Studies concluded that this increased trend in disease prevalence is accurate and does not reflect an upward bias due to Coal Workers’ Health Surveillance Program (CWHSP) surveillance efforts (Laney and Attfield 2014). Although there is no clear explanation for the increase in CWP cases, theories include miners drilling through rock to access thinner coal seams, longer working hours leading to increased coal dust exposure, and lack of compliance with safety regulations. These theories are discussed further in Section 3.3.1.

[17] Summit’s scan of datasets for use in quantitative analysis found no available data prior to 1970. The Federal Mine Safety and Health Act of 1977 amended the Coal Mine Safety and Health Act of 1969.

Figure 4: Nationwide prevalence of coal workers' pneumoconiosis among coal miners with 25+ years of experience, from literature review findings



Source: Blackley, Halldin, and Laney 2018.

CWP represents the narrowest definition of black lung disease. However, within government sources, the definition of black lung disease varies. For example, although many government sources (such as the Black Lung Benefits Act) borrow the definition of CWP to define black lung disease, a definition from an archived MSHA website titled *End Black Lung: Act Now* (accessed January 24, 2023) defines black lung disease as CWP, emphysema, silicosis, and bronchitis.

Coal mine dust lung disease (CMDLD). The term CMDLD has been introduced to capture a broader spectrum of disease associated with exposure to coal dust, which can include silicosis (a type of pneumoconiosis caused by inhalation of silica dust), CWP, mixed-dust pneumoconiosis (associated with exposure to both coal and crystalline silica dusts), PMF, dust-related diffuse fibrosis, COPD, and chronic airway diseases, such as emphysema and chronic bronchitis (Karatela, Caruana, and Paul 2022; Somers 2017; Laney and Weissman 2014; Petsonk, Rose, and Cohen 2013). This category is shown in the second row in Figure 3.

Other respiratory illnesses. Exposure to coal through coal mining, residential coal use, or outdoor air pollution can also increase rates of other respiratory illnesses not typically included in the definitions of black lung disease or CMDLD. These other respiratory illnesses include lung cancer, asthma (particularly in young children), decreased lung function, and acute lower respiratory infections (particularly in children under 5) (Finkelman, Wolfe, and Hendryx 2021; Kerimray et al. 2017; Buchanan, Burt, and Orris 2014; Petsonk, Rose, and Cohen 2013; Hosgood et al. 2010; Galeone et al. 2008; Torres-Duque et al. 2008). This category is shown in the third row in Figure 3.

Broader health impacts. Negative health impacts associated with coal exposure beyond respiratory illnesses include low birth weight in newborn infants, increased infant mortality, neurological effects, mental illness, cataracts, immune system impairment, cardiovascular problems, and chronic heart, lung, and kidney diseases (Karatela, Caruana, and Paul 2022; Finkelman, Wolfe, and Hendryx 2021; Zierold, Hagemeyer, and Sears 2020; Braithwaite et al. 2019; Kerimray et al. 2017). This category of health impacts is shown at the bottom of the triangle in Figure 3.

3.1.2. Non-occupational exposure to coal dust

While much of the research on black lung disease is focused on occupational exposure through coal mining or processing, exposure to coal dust through residential use or air pollution is also associated with black lung disease and other respiratory illnesses.

> Coal mining is not the only avenue for negative health impacts; residential coal burning and coal-related outdoor air pollution can also lead to negative health impacts.

Residential coal use. Extensive reviews of publications on the health effects of residential coal use have found a statistically significant correlation between residential solid fuel use and risk of lung cancer, acute respiratory infection, and COPD (Hosgood et al. 2011; Zhang and Smith 2007). For example, decades of scientific research among nonsmoking women in Xuanwei, China, has shown their uncommonly high rates of lung cancer are associated with their reliance on coal for heating and cooking and maintenance of open ash pits used in agriculture (Barone-Adesi et al. 2012; Large et al. 2009; Zhang and Smith 2007). Research also indicates that children are particularly susceptible to coal-related illness due to residential coal exposure; in 2020, the World Health Organization estimated that the inhalation of particulate matter such as soot from household air pollution is responsible for almost half of all fatal lower respiratory infections among children under 5 years of age (World Health Organization 2022).

Outdoor air pollution. Studies have also shown that coal dust produced by mining, coal transportation, and processing can disperse to nearby communities and impact local air pollution (Huertas et al. 2014; Huertas et al. 2012; Mandal et al. 2012). A systematic review of ecological studies found that villages closer to coal mines had higher risks of mortality and morbidity from diseases including circulatory and respiratory diseases, congenital abnormalities, and cancer (Cortes-Ramirez 2018).

3.1.3. Diagnosing black lung disease and measuring incidence

Diagnosing black lung disease. Diagnosis of respiratory disease due to coal mine dust exposure requires respiratory symptoms, medical tests such as lung imaging and pulmonary function testing, and a detailed history of exposure (often occupational) (Petsonk, Rose, and Cohen 2013). Because the symptoms and respiratory illnesses associated with exposure to coal dust can easily resemble other respiratory illnesses, the only way to conclusively pinpoint the cause to coal dust is through a well-documented history of exposure (Arnold 2016) or through autopsies. This makes the history of exposure "a critical component in the diagnosis" (Petsonk, Rose, and Cohen 2013).

International Labour Organization standards and B Readers. Extensive documentation exists providing detailed instruction on interpreting the specific size, location, and appearance of lesions and other symptoms to diagnose black lung disease. The International Labour Organization (ILO) publishes standards for classifying pneumoconiosis to ensure uniformity in the description and recording of "radiographic abnormalities in the chest provoked by the inhalation of dusts" (ILO 2011). To ensure that the ILO classification scheme is applied consistently, NIOSH created the B Reader Program in 1974 to teach physicians the ILO classification system and certify them as "B readers" upon completion of their training. (Chest Radiography 2022; Zosky et al. 2016). For a coal worker to be eligible for compensation under the Federal Black Lung Program, NIOSH requires two certified B Readers to classify radiographs "for the presence, profusion, and type of lung parenchymal abnormalities" (Blackley, Halldin, and Laney 2018).

Medical diagnostic codes. Table 8 identifies the relevant black lung disease diagnostic codes from the International Classification of Diseases (ICD) classification system (CMS 2022).[18] With the exception of the final row, all ICD codes in the table below require exposure for diagnosis; for example, ICD-10 CM code J62 requires pneumoconiosis to be directly attributable to silica exposure. There may be instances where a healthcare professional is unable to make a causal connection to talc dust, silica, or other dust due to lack of medical records or awareness of a patient's occupational history (Kurth and Casey 2020). In this situation, the healthcare professional may use other ICD codes related to respiratory and cardiovascular diagnoses, but not specifically associated with talc dust, silica, or other dust; this can result in a misclassification of disease (Kurth and Casey 2020).

Table 8: Black lung disease diagnostic codes

Diagnosis Description*	ICD-9 CM Code**	ICD-10 CM Code***
Coal Workers' Pneumoconiosis (CWP)	500	J60
Asbestosis (Pneumoconiosis due to asbestos and other mineral fibers)	501	J61
Pneumoconiosis due to dust containing silica****	-	J62
Pneumoconiosis due to other dust containing silica	502	J62.8
Berylliosis, pneumoconiosis due to other inorganic dust	503	J63.2
Pneumonopathy due to inhalation of other dust	504	-
Pneumoconiosis, unspecified	505	J64

*Table sources: CMS 2022, Kurth and Casey 2020
**International Classification of Diseases (ICD) Clinical Modification (CD) is a diagnostic system used to code and classify medical diagnoses.
*** The universal use of ICD-10 codes was mandated in 2015 (CMS 2021), but the preceding ICD-9 codes may also be relevant in studies of historical black lung disease prevalence.
****Also known as silicosis

Overall, the diagnosis of black lung disease is a complex process, relying on certified B Readers following ILO standards to examine specific size, location, and appearance of lesions in addition to other common symptoms, such as a chronic cough or shortness of breath. Because the symptoms and respiratory illnesses associated with coal dust can often resemble other illnesses, a well-documented history of exposure or an autopsy is the only conclusive way to determine the cause (Arnold 2016; Petsonk, Rose, and Cohen 2013). However, even with sophisticated imaging technology following the ILO standards, there is evidence of underdiagnosis of black lung disease, as discussed below.

Challenges in measuring black lung incidence. There are several federal programs that aim to measure black lung incidence across the United States. For example, MSHA's mine dataset gathers data quarterly from mine operators, including illnesses reported by mine operators and contractors through MSHA Form 7000-1 (MSHA, "Accident Injuries Data Set"). Separately, under the NIOSH CWHSP, NIOSH staff travel across the United States conducting mobile black lung disease screenings through chest X-rays and spirometry tests with miners. However, it is difficult to measure the full universe of black lung

[18] The universal use of ICD-10 codes was mandated in 2015 (CMS 2021), but the preceding ICD-9 codes may also be relevant in studies of historical black lung disease prevalence. Not all ICD-10 codes for lung diseases in the range between J60 and J70 were relevant for analysis; those outside the range of acceptable diagnoses based on the literature were excluded from this study. For example, cannabinosis (J66.2), is a disease stemming from routine exposure to cannabinoids (i.e., marijuana use), which is not relevant to the purpose of this study.

incidence in the United States, as there are measurement challenges associated with undiagnosed cases, diagnosed cases, and deaths as described below.

Black lung deaths. It is easiest to measure black lung incidence based on deaths because black lung disease can be definitively diagnosed through an autopsy. However, a determination of black lung incidence in the United States based on deaths is naturally conservative because it will miss cases (diagnosed or undiagnosed) of people living with black lung disease. For example, a miner suffering from pneumoconiosis who has never received an X-ray for a formal diagnosis or a coal plant worker whose lung cancer was attributed to smoking would be missed. In order to protect workers' privacy, WONDER only reports results where 10 or more cases were diagnosed between 1999 and 2020. Counties with fewer than 10 cases appear as zero values with no way to discern these counties from those with zero diagnosed cases.

Diagnosed black lung cases. Measuring incidence based on diagnosed cases in addition to deaths is more comprehensive but complicated by the definitional ambiguity illustrated in Section 3.1.1. The count of diagnosed black lung cases will change depending on how broad (all respiratory diseases associated with coal dust inhalation) or narrow (only CWP) a definition of black lung disease is used in data collection. ECWHSP is a publicly funded project. In order to protect workers' privacy, ECWHSP only reports results where 10 or more cases were diagnosed between 1970 and 2014. Counties with fewer than 10 cases appear as zero values with no way to discern these counties from those with zero diagnosed cases.

Undiagnosed black lung cases. Undiagnosed cases pose a large challenge for accurately measuring black lung incidence in the United States, given undercounts in diagnosed cases. Blackley, Halldin, and Laney (2018) and Shriver and Bodenhamer (2018) have written about the reasons for likely undercounts among miners, as presented below.

First, there is low uptake of the voluntary, free radiograph screenings that the CWHSP offers to active and former coal miners. Although these screenings are crucial in identifying CWP and determining disease prevalence, the CWHSP[19] has an annual participation rate of only 30% to 40% of active miners (Potera 2019; Blackley, Halldin, and Laney 2018). Shriver and Bodenhamer (2018) theorize several reasons for the low participation rate, including a disincentive to seek detection or diagnosis opportunities due to the risk of being fired and logistic difficulties in traveling to distant screening facilities.[20] Additionally, miners have reported fears of workplace retaliation upon seeking diagnoses (Shriver and Bodenhamer 2018). In a 2016 *PBS NewsHour* interview, National Public Radio's (NPR's) Howard Berkes explained that although it is illegal for mining companies to fire miners due to a black lung disease diagnosis, "every single miner I have talked to in Appalachia in the last 6 years has said the same thing.… What they fear is just even going to the NIOSH vans that come into their communities with X-ray equipment and being seen going into those vans—just that—can cause the mining company to say 'This guy might have black lung'" (*PBS NewsHour* 2016). The miners' opinion was that "if the mining company finds out, they'll lose their jobs, so they don't get tested" (*PBS NewsHour* 2016).

[19] The CWHSP conducts free chest radiographs via a mobile van for miners and has been collecting data since the program's inception in 1970 (Centers for Disease Control 2020).
[20] Some states like Kentucky have a statute of limitations requiring miners to make a workers' compensation claim within a certain time frame, which could make it more difficult to successfully file a claim (Shriver and Bodenhamer 2018).

While quantitative data can be limited to specific ICD codes that encompass black lung, literature review sources were less consistent. Some sources limited their research to only CWP, others used the term to refer to a broader set of diseases (along the lines of CMDLD), and still others did not define black lung disease within the context of their paper.

Additionally, low worker's compensation approval rates may lead to undercounting of cases. The rate of miners successfully receiving a CWP diagnosis and approval for compensation through the Federal Black Lung Program is relatively low, although it has increased since the 1990s. In the 1990s, only 4% of initial benefit applications were approved (Toler 2002), compared to 15% of claims in 2013 (Hamby, Ross, and Mosk 2013), and 32% in 2022 (DOL 2023). The appeals process for denied claims is often lengthy, expensive, and unsuccessful for many miners (Royal 2019, Cartwright 2016). Cartwright (2016) points out that by the time a miner's case is won on appeal, coal mining companies have gathered evidence to reappeal and win based on a radiograph that has been reexamined by a different expert who has come to the opposite conclusion. In a letter to the editor of the *Social Determinants of Health* journal, an academic descended from Appalachian coal miners reported that in his experience miners feel that the Division of Coal Mine Workers' Compensation request system is "unjust and biased against them" (Royal 2019). As Royal (2019) describes, low approval rates and the complex appeals process—along with other challenges such as the belief that Black Lung Program claims are prohibitively expense to file—"prevent many miners from pursuing black lung cases."

3.1.4. Black lung definition used for this study

The black lung definition used in this study was determined by insights gathered during the literature review coupled with publicly available data. As described above, the quantitative data used in this study are drawn from ECWHSP (black lung cases) and WONDER (black lung deaths). ECWHSP defines and tracks black lung cases according to the single ICD-10 diagnosis code J60, which indicates CWP. This means that other diagnoses that can be attributed to coal exposure (such as silicosis, ICD-10 code J62) are not captured in these data. However, the team was able to analyze black lung deaths according to its expanded definition of black lung disease via the CDC's WONDER database. WONDER tracks all deaths according to relevant diagnosis code, so the team opted to review the data according to the list of black lung diagnoses in Table 8 to capture a more expansive view of black lung disease than the strict CWP tracking done by ECWHSP.

3.2. Black lung disease prevalence

The research team generated summary statistics of black lung cases and deaths overall. Given that each metric is defined and collected differently as discussed in Section 2.2.1 and 2.2.2, Summit analyzed black lung cases and deaths independently.

3.2.1. Black lung cases

The team found that 103 counties (3.3%) reported 10 or more black lung cases between 1970 and 2014.[21] Of these, 132 cases were reported per county on average across the 44-year period. These cases include all three stages of the ILO measurements as well as PMF cases—the most severe form of black

[21] Recall that ECWHSP only reports county-level results where 10 or more cumulative black lung cases were found. Additionally, ECWHSP defines and tracks black lung disease according to the single ICD-10 diagnosis code J60, known as coal workers' pneumoconiosis. This differs from how we measure black lung deaths, discussed in the next section.

lung disease.[22] Keeping in mind that data is only suppressed if the county has less than 10 total black lung cases during the collection period, 30 of counties report at least one PMF case. Cambria County, PA had 105 reported PMF cases between 1970 and 2014.

Table 9: Cumulative black lung cases by severity across the entire collection period, 1970-2014

Severity Level	# of Counties with 1+ case	Min. # cases per county	Median # cases per county	Avg. # cases per county	Max. # cases per county
ILO Level 1 (least severe)	103	10	48	100.00	562
ILO Level 2	50	0	0	22.66	183
ILO Level 3	5	0	0	0.72	26
PMF (most severe)	30	0	0	8.61	105
OVERALL	**103**	**10**	**54**	**131.99**	**876**

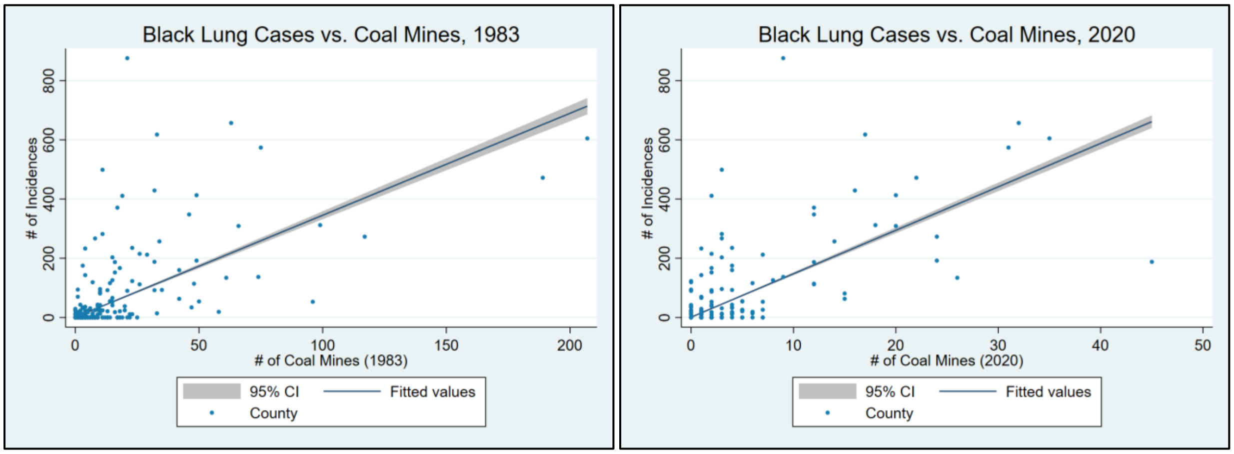
Source: Enhanced Coal Workers' Health Surveillance Program (ECWHSP). Counties with 10 or more total diagnosed cases at any level are reported by ECWHSP. Note: ILO stands for International Labour Organization.

Of these 103 counties with black lung cases, all of them have at least one case stemming from underground mines. Only 16 counties report black lung cases from surface mines, none of which exceeded an ILO rating of 1, the least severe form of black lung disease.

The team compared the number of number black lung cases to the various metrics—namely number of coal mines, short tons of coal produced, number of employees, exposure hours, number of households using coal or coke. Summit used correlation matrixes to understand the relationship between black lung cases and these metrics across each data collection period.[23] Of these factors, total weekly coal dust exposure (1970–2020) has the highest correlation with black lung cases (0.7390); 2020 coal production has the lowest (0.0850). Additionally, for a visual representation of the data, Summit compared black lung cases to these factors using scatterplots, shown in Figure 5 through Figure 9 below. The x-axis presents the relevant coal metric and year of data collection (e.g., number of coal mines in 1983) while the y-axis depicts the number of black lung cases. Each dot in the graph represents one county. The diagonal line in the graph represents the average county for reach combination of coal mines and black lung cases at 95% confidence.

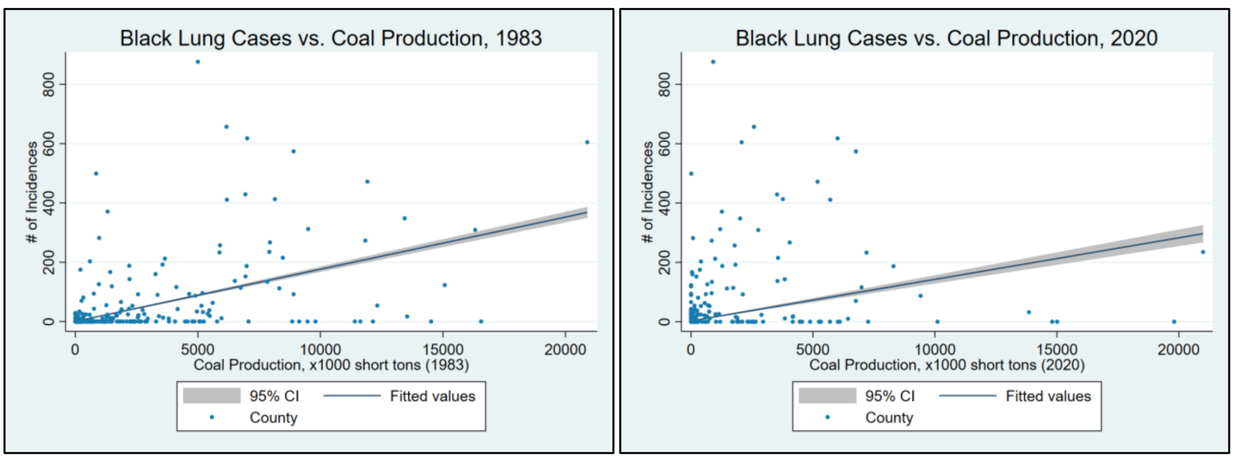
[22] The stages of black lung disease are specified in Section 2.2.2.
[23] Correlation matrixes can be found in Appendix B. Additional details on data collection periods can be found in Section 2.2.1.

Figure 5: Cumulative black lung cases versus number of coal mines (1983, 2020)



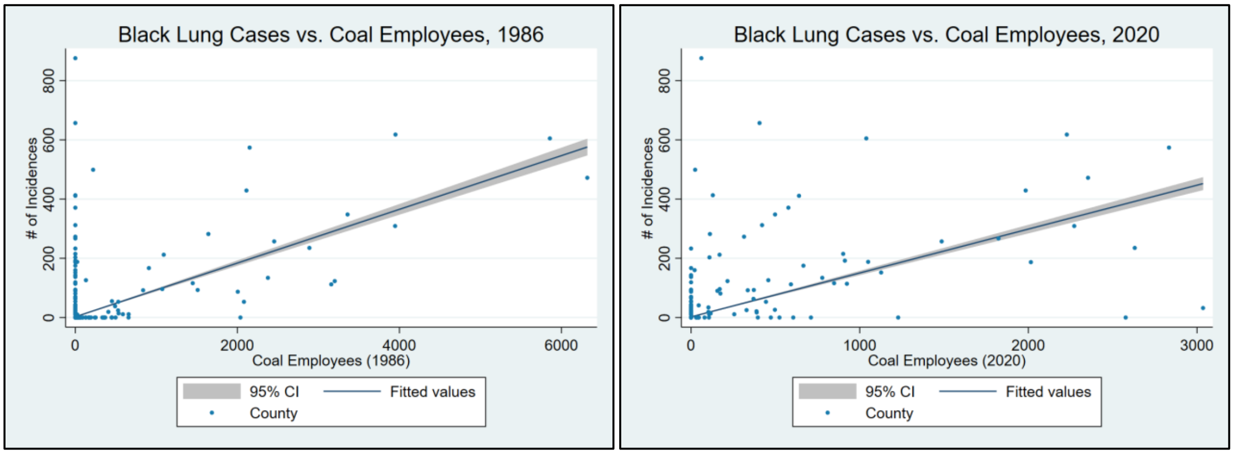
Sources: Center for Disease Control and Prevention (CDC) Enhanced Coal Workers' Health Surveillance Program (ECWHSP) and U.S. Energy Information Administration (EIA). Note: 1983 is the earliest coal mine data available from EIA; 2020 is the most currently available. Data from ECWHSP overlaps with both time periods (1970-2014). One outlier observation dropped to enhance the graphic, totaling 3,135 counties for analysis.

Figure 6: Cumulative black lung cases versus amount of coal production in short tons (1983, 2020)



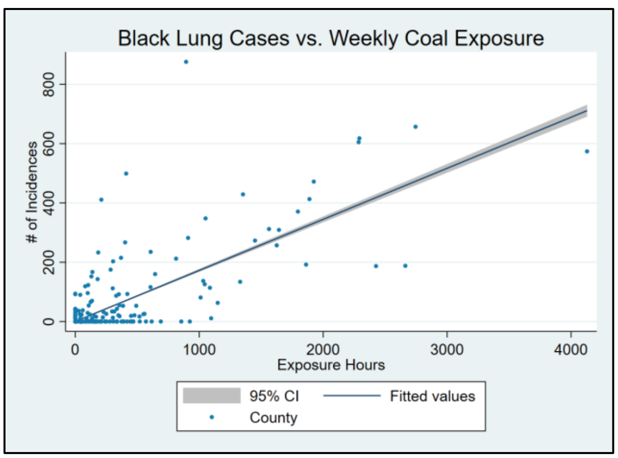
Sources: Center for Disease Control and Prevention (CDC) Enhanced Coal Workers' Health Surveillance Program (ECWHSP) and U.S. Energy Information Administration (EIA). Note: 1983 is the earliest coal mine data available from EIA; 2020 is the most currently available. Data from ECWHSP overlaps with both time periods (1970-2014). One outlier observation dropped to enhance the graphic, totaling 3,135 counties for analysis.

Figure 7: Cumulative black lung cases versus number of coal employees (1986, 2020)



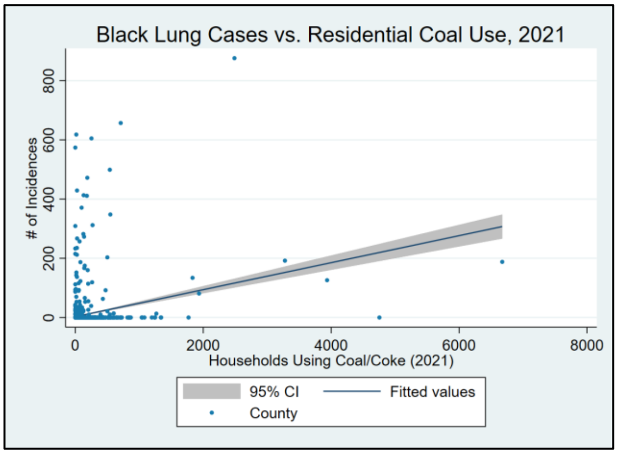
Sources: Center for Disease Control and Prevention (CDC) Enhanced Coal Workers' Health Surveillance Program (ECWHSP) and U.S. Census County Business Patterns (CBP). Note: 1986 is the earliest coal mine data available from CBP; 2020 is the most currently available. Data from ECWHSP overlaps with both time periods (1970-2014). One outlier observation dropped to enhance the graphic, totaling 3,135 counties for analysis.

Figure 8: Cumulative black lung cases versus weekly hours of coal dust exposure (hours, 1970–2020)



Sources: Center for Disease Control and Prevention (CDC) Enhanced Coal Workers' Health Surveillance Program (ECWHSP) and Mine Safety and Health Administration (MSHA) Mines Data Set. Note: Data are collected quarterly from 1970-2020. 3,136 counties represented for analysis. X-axis depicts the total number of exposure hours across all mines, with 24 possible hours of exposure per mine. One outlier observation dropped to enhance the graphic, totaling 3,135 counties for analysis.

Figure 9: Cumulative black lung cases versus number of households burning coal (2021)



Sources: Center for Disease Control and Prevention (CDC) Enhanced Coal Workers' Health Surveillance Program (ECWHSP) and U.S. Census Bureau American Community Survey (ACS) 1-year estimates. Note: Residential coal use data sourced from 2021 U.S. Census American Community Survey. One outlier observation dropped to enhance the graphic, totaling 3,135 counties for analysis.

All of the results up to this point are raw counts, meaning that they do not take county size into account. Factoring a county's size into the results, meanwhile, is called the black lung case rate. If, for example, counties with very few residents have a substantially higher rate of black lung cases relative to their size, then the real prevalence of black lung disease in that county is significantly more pervasive than counties who have the same number of cases but a larger population. When the research team took each county's population into account, it became clear that some counties are far more susceptible to black lung diagnoses than others. United States counties average 4.34 black lung cases between 1970 and 2014. However, the black lung rate is roughly is roughly 0.11 cases per 1,000 residents—or 11 cases per 100,000—using 2014 population statistics.

> ***Research Question 1.1:*** What is the total number and rate per 1,000 residents of black lung cases?
>
> **Answer: 0.11 black lung cases, or 11 cases per 100,000 residents.**

This comparison becomes particularly illuminating when you compare the top counties across black lung case counts to case rates. For example, McDowell County, West Virginia—population 20,385 in 2014—ranks second overall in terms of count of black lung cases (657 cases). Cambria County, Pennsylvania, whose population exceeds 137,000, is the only county with more black lung cases (876). When population size is taken into account, McDowell becomes the country's densest county in terms of the rate of black lung cases (32.23 cases per 1,000 residents), while Cambria falls in rank to 16th (6.37 cases per 1,000). Table 10 below illustrates the trade-off of the top 10 counties relative to black lung counts and rates.

Table 10: Top ten counties in terms of cumulative black lung cases and rates, 1970-2014

Rank	County	Number of Cases (Counts)	County	Number of Cases (Rate per 1,000)
Mean U.S. County		***4.34***		***0.11***
1	Cambria, PA	876	McDowell, WV	32.23
2	McDowell, WV	657	Buchanan, VA	20.39
3	Raleigh, WV	618	Wyoming, WV	18.29

Rank	County	Number of Cases (Counts)	County	Number of Cases (Rate per 1,000)
Mean U.S. County		***4.34***		***0.11***
4	Pike, KY	605	Logan, WV	16.28
5	Logan, WV	574	Boone, WV	13.03
6	Fayette, PA	499	Harlan, KY	11.14
7	Buchanan, VA	472	Mingo, WV	9.98
8	Kanawha, WV	429	Pike, KY	9.61
9	Wyoming, WV	413	Dickenson, VA	8.93
10	Washington, PA	411	Nicholas, WV	8.24

Sources: Enhanced Coal Workers' Health Surveillance Program (ECWHSP) and U.S. Census Bureau American Community Survey (ACS) 1-year estimates.

3.2.2. Black lung deaths

3.2.2.1. Raw counts and rates

The research team created and reviewed a similar set of summary statistics for black lung deaths. Namely, the team reviewed the black lung death counts and rates across counties to better understand how geological and geographic factors correlate with black lung death statistics. The key difference between black lung case data and death data, however, is the flexibility in defining black lung disease. The ECWHSP, responsible for monitoring live black lung cases, maintains a strict definition of black lung disease—namely CWP (ICD-10 code J60). However, using information gathered through literature and expert reviews, the research team expanded its definition of black lung disease when reviewing death data via the WONDER database (discussed in Section 3.1.3). This allows the team to compare NIOSH's strict interpretation of black lung disease against expanded and additional interpretations of black lung disease, including silicosis. While some recorded deaths from other diagnosis codes may not be directly attributable to coal dust exposure (e.g., asbestosis), the marginal cost of overcounting is expected to be low compared to the risk of undercounting black lung deaths by only referring to CWP.

Upon review, 333 counties report 10 or more black lung deaths according to the team's expanded definition. By contrast, only 65 counties report CWP deaths—the stricter definition used to track cases of living individuals by ECWHSP. The combination of the CWP diagnosis (J60) along with others, such as silicosis, provides a more nuanced view of deaths attributable to black lung disease. Under this expanded definition of black lung, asbestosis (J61) provides the most likely source of overcounting deaths. Asbestosis cases can be found in 277 counties during the WONDER database's collection period (1999-2020), most of which are not coal counties. These cases are common in coastal areas such as California, Florida, and New Jersey, suggesting that many of these deaths are not among coal workers.

Table 11 below provides high-level summary statistics of the research team's total definition for black lung disease as well as diagnosis code–level results for a more granular view.

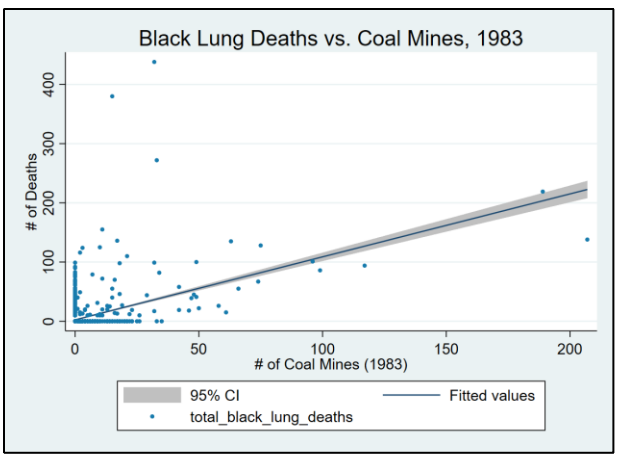
Table 11: Summary statistics of cumulative black lung deaths by diagnosis code across the entire collection period, 1999-2020

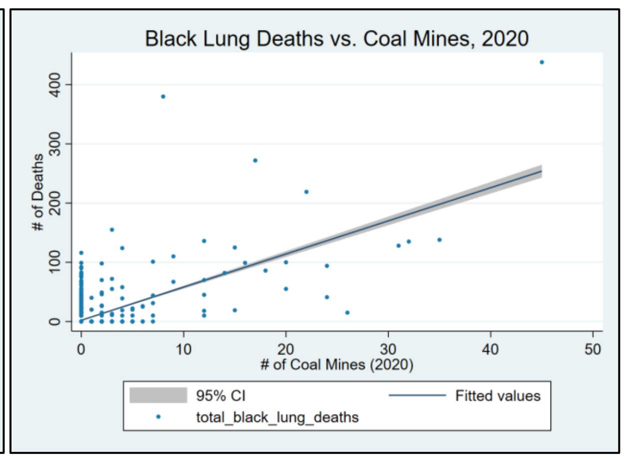
Diagnosis	# of counties with 10+ deaths	Min. # deaths per county	Median # deaths per county	Avg. # deaths per county	Max. # deaths per county
Coal workers' pneumoconiosis	65	0	0	10.67	401
Asbestosis (Pneumoconiosis due to asbestos and other mineral fibers)	277	0	15	19.19	99
Pneumoconiosis due to talc dust	0	0	0	0	0
Pneumoconiosis due to other silica or silicates (silicosis)	10	0	0	0.38	20
Pneumoconiosis due to other inorganic dust	1	10	10	10.00	10
Pneumonopathy due to inhalation of other dust	0	0	0	0	0
Pneumoconiosis, unspecified	31	0	0	2.10	61
TOTAL	**333**	**10**	**19**	**32.37**	**438**

Source: Center for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) database.

The research team compared black lung deaths to various coal metrics. Reviewing black lung deaths across number of coal mines, short tons of coal produced, number of employees, exposure hours, number of households using coal/coke, the team observed a similar set of correlations as it did when analyzing black lung cases. Scatterplots comparing black lung deaths to coal-related metrics can be found in the Figures below.

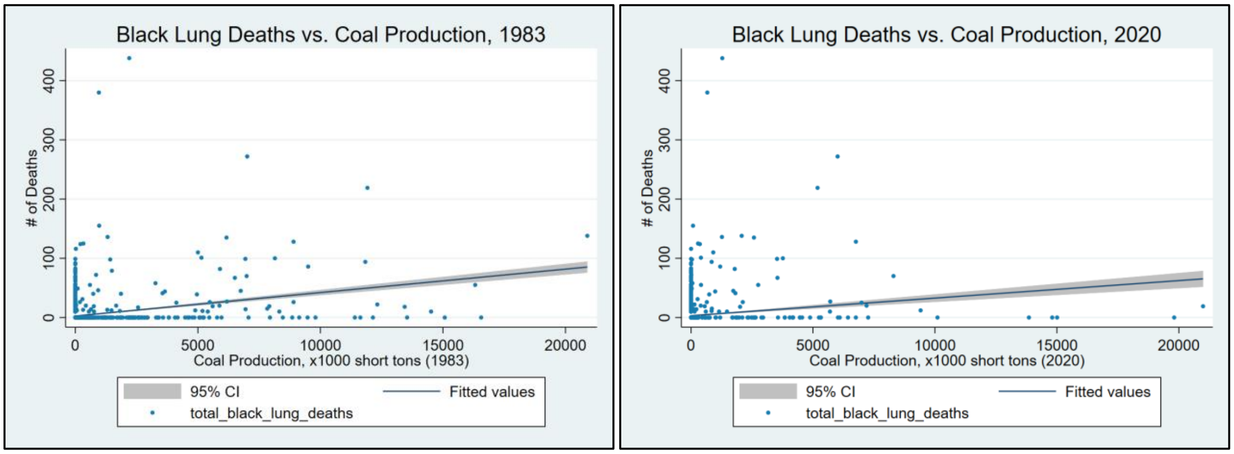
Figure 10: Cumulative black lung deaths versus number of coal mines (1983, 2020)





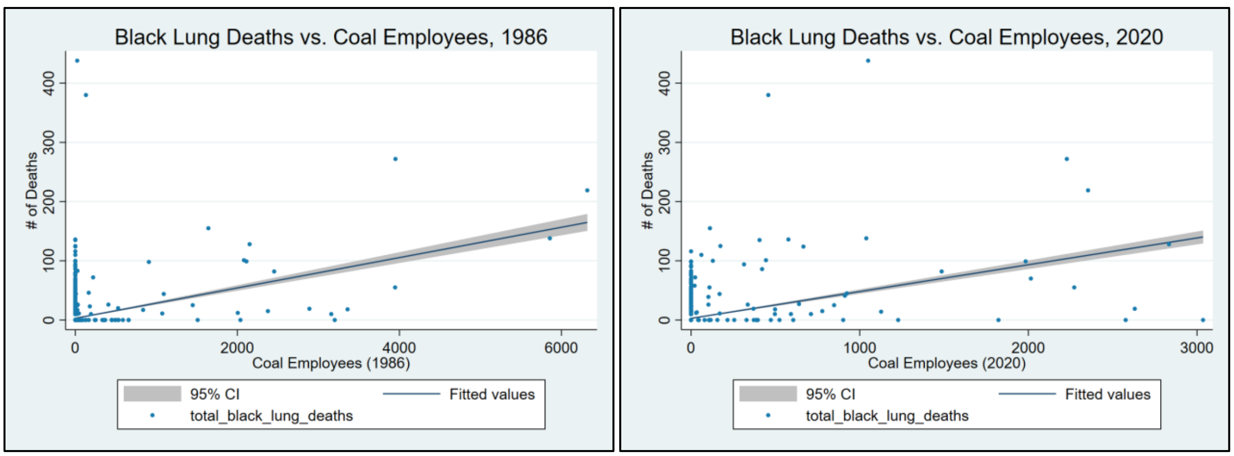
Sources: Center for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) database and U.S. Energy Information Administration (EIA). Note: 1983 is the earliest coal mine data available from EIA; 2020 is the most currently available. Data from CDC WONDER spans 1999-2020. One outlier observation dropped to enhance the graphic, totaling 3,135 counties for analysis.

Figure 11: Cumulative black lung deaths versus amount of coal production in short tons (1983, 2020)



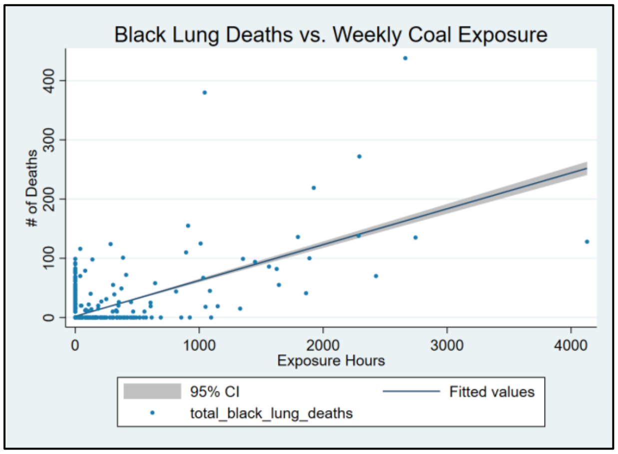
Sources: Center for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) database and U.S. Energy Information Administration (EIA). Note: 1983 is the earliest coal mine data available from EIA; 2020 is the most currently available. Data from CDC WONDER spans 1999-2020. One outlier observation dropped to enhance the graphic, totaling 3,135 counties for analysis.

Figure 12: Cumulative black lung deaths versus number of coal employees (1986, 2020)



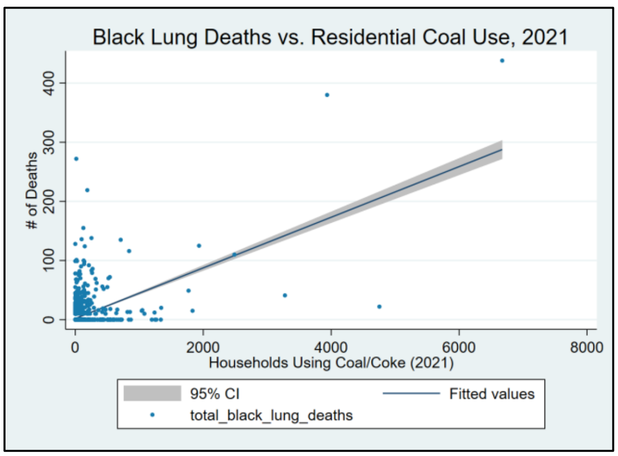
Sources: Center for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) database and U.S. Census County Business Patterns (CBP). Note: 1986 is the earliest coal mine data available from CBP; 2020 is the most currently available. Data from CDC WONDER spans 1999-2020. One outlier observation dropped to enhance the graphic, totaling 3,135 counties for analysis.

Figure 13: Cumulative black lung deaths versus weekly hours of coal dust exposure (hours, 1970–2020)



Sources: Center for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) database and Mine Safety and Health Administration (MSHA) Mines Data Set. Note: Data are collected quarterly from 1970-2020. 3,136 counties represented for analysis. X-axis depicts the total number of exposure hours across all mines, with 24 possible hours of exposure per mine. One outlier observation dropped to enhance the graphic, totaling 3,135 counties for analysis.

Figure 14: Cumulative black lung deaths versus number of households burning coal (2021)

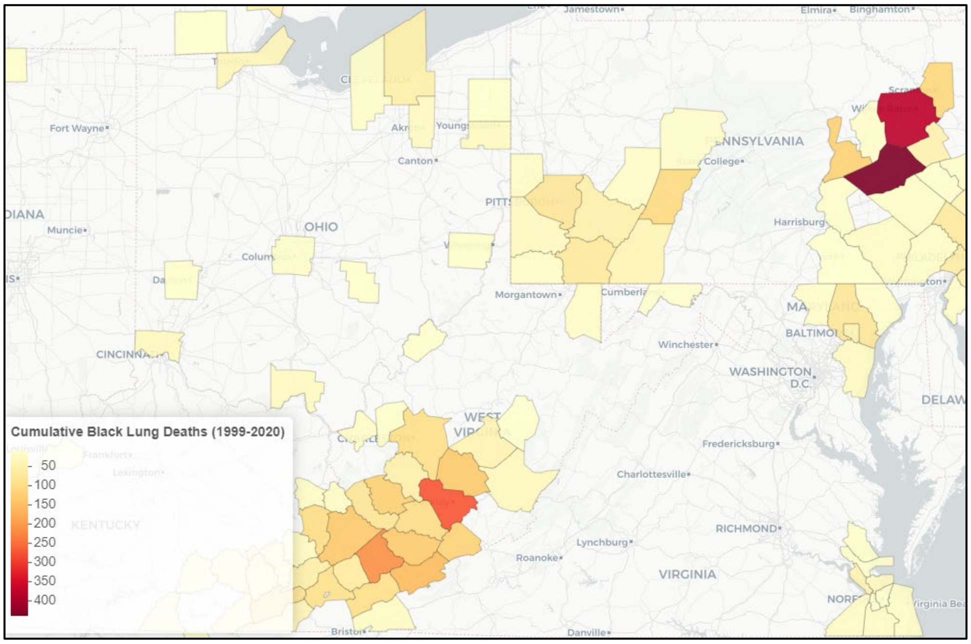


Sources: Center for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) database and U.S. Census Bureau American Community Survey (ACS) 1-year estimates. Note: Residential coal use data sourced from 2021 U.S. Census American Community Survey. One outlier observation dropped to enhance the graphic, totaling 3,135 counties for analysis.

In addition to the raw count black lung deaths, it is important to understand the mortality rate, meaning relative to county population. While the average county across the entire United States has 3.44 black lung deaths between 1999 and 2020, the death rate is roughly 0.04 deaths per 1,000 residents in 2020, or four deaths per 100,000 according to 2020 Census data. Figure 15 and Figure 16 exemplify the difference between overall black lung deaths and death rates controlled for county population size using counties in Northern and Central Appalachia (parts of Pennsylvania, Maryland, and Virginia, West Viginia, and Ohio) as an example. The counties shaded darker represent higher tallies.

> ***Research Question 1.2:*** What is the total number and rate per 1,000 residents of black lung deaths?
>
> **Answer: 0.04 black lung deaths, or 4 deaths per 100,000 residents.**

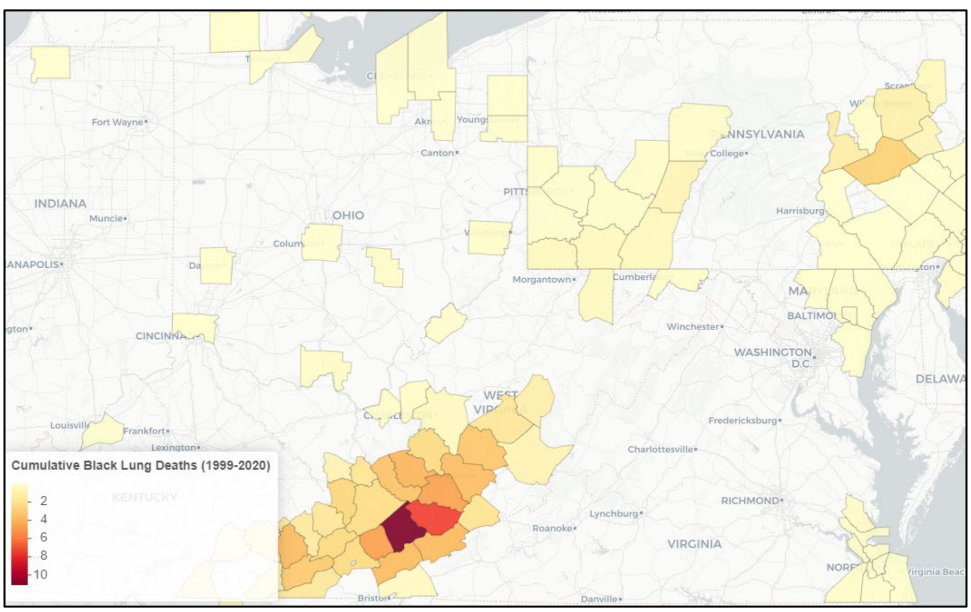
Figure 15: Cumulative black lung deaths (1999–2020)



Heat map of county-level results for Ohio, Viginia, West Virginia, Maryland, Delaware, and Pennsylvania. Results show the cumulative number of black lung deaths between 1999 and 2020. Reports ranges from 10 cumulative deaths to over 400 for a single county.

Source: Center for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) database.

Figure 16: Cumulative black lung death rates (per 1,000 residents in 2020) (1999–2020)



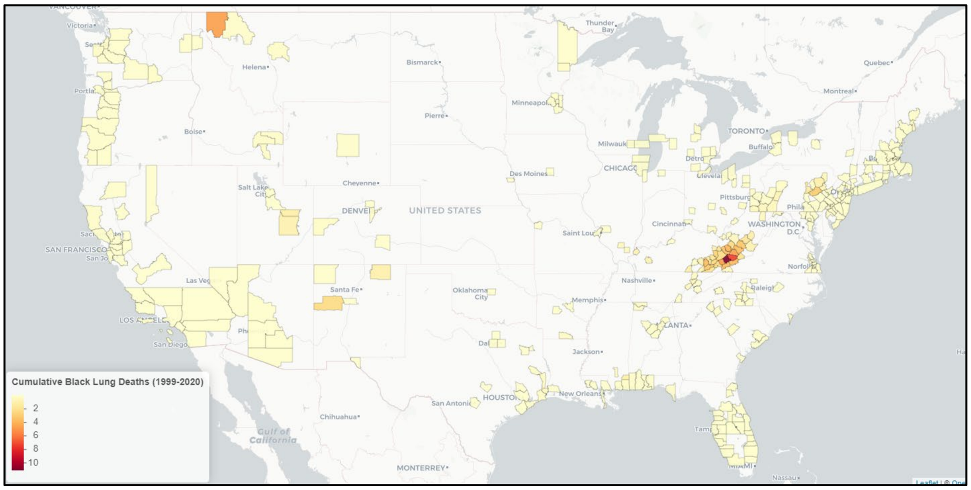
Heat map of county-level results for Ohio, Viginia, West Virginia, Maryland, Delaware, and Pennsylvania. Results show the cumulative number of black lung deaths per 1,000 residents between 1999 and 2020. Reports ranges from less than 1 death per 1,000 residents to more than 10 deaths per 1,000.

Sources: Center for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) database and U.S. Census Bureau American Community Survey (ACS) 1-year estimates.

3.2.2.2. Comparing black lung definitions

Under the research team’s expanded definition of black lung using 7 total ICD-10 diagnosis codes, the death rate per 1,000 residents in 2020 is 0.04 deaths, or 4 deaths per 100,000. If, however, we compare the death rate to deaths only from CWP (J60), the death rate drops to 0.03 deaths per 1,000 (3 deaths per 100,000). While the current definition helps us capture tangential black lung deaths not classified as CWP, there is a chance for some overcounting, particularly among asbestosis (J61) deaths as discussed in Section 3.2.2.1. Figure 17 and Figure 18 below show the difference across the United States when black lung is measured across all seven diagnosis codes versus CWP only.

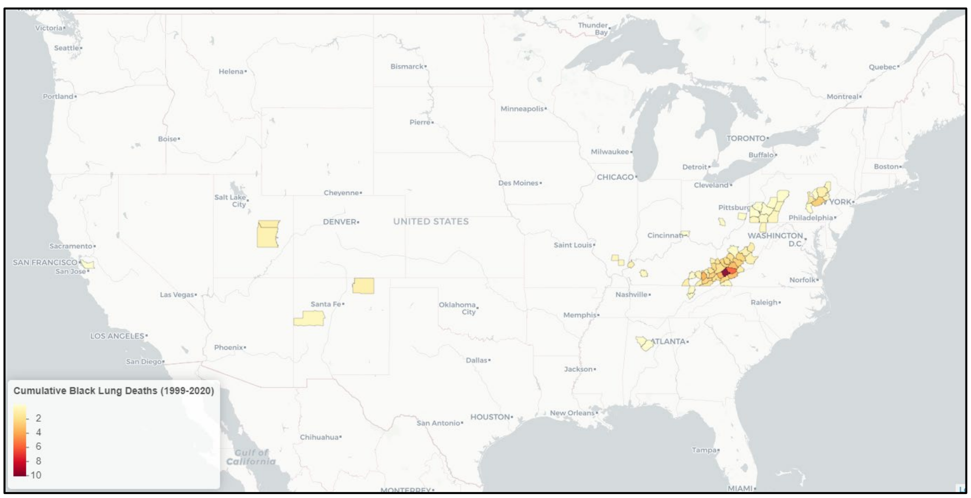
Figure 17: Cumulative black lung death rates among all ICD-10 diagnoses (1999–2020)



Heat map of the continental United States displaying the total number of black lung deaths across all black lung ICD-10 diagnoses. Appalachia maintains the densest concentration of cumulative deaths, with clusters of other counties in coastal areas including Florida, California, Delaware, and New York.

Source: Center for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) database.

Figure 18: Cumulative black lung death rates attributable to CWP (1999–2020)



Heat map of the continental United States displaying the total number of black lung deaths attributable to CWP only. Appalachia maintains the densest concentration of cumulative deaths. Coastal areas including Florida and California yield zero results, contrasting the results in Figure 17 when the definition of black lung is restricted to remove other diagnoses, particularly asbestosis.

Source: Center for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) database.

3.3. Black lung prevalence across county types and populations of interest

In this section, we compare black lung prevalence across county types (coal counties, former coal counties, non-coal counties, residential use counties) and populations of interest (Appalachia and the Navajo Nation). This section begins with contextual information on both populations of interest, before delving into the results of statistical comparisons of prevalence.

3.3.1. Context on Appalachian populations

This section provides contextual information from the literature review on coal mining and residential coal use in Appalachian populations.

Appalachian coal mines. Appalachia is experiencing a resurgence of CWP cases, beginning with a sharp rise after the 1990s (Blackley, Halldin, and Laney 2018). Based on data collected between 2013 and 2017, one in five Central Appalachian coal miners with over 25 years of experience has evidence of CWP (a diagnosis which may include PMF and silicosis), which is "the highest level recorded during the past 25 years" (Blackley, Halldin, and Laney 2018). There are several theories to explain the increase in CWP cases among Appalachian coal miners:

- *Drilling through rock to access thinner coal seams.* Because many of the easily accessible and purer coal seams have already been mined, some of the remaining coal in Appalachia is in thinner or narrower seams buried beneath layers of silica-rich rock. The emergence of more powerful machines allows miners to drill through this rock, increasing their exposure to mixed-dust particulate matter and increasing silica-related lung health risks (Sisk 2023; Hamby 2020, 275; Ranavaya II, Ranavaya, and Chongswatdi 2020; Shriver and Bodenhamer 2018; Somers 2017; Cartwright 2016).

- *Long working hours.* Present-day miners are working longer hours than they did prior to the resurgence of black lung disease (Hamby 2020, 274; Arnold 2016; Cartwright 2016). Researchers theorize that longer hours increase overall exposure and decrease physical recovery time between mining shifts, potentially leading to higher rates of CWP (Hamby 2020, 274).

- *Lack of compliance with safety regulations.* There is evidence of systematic disregard of legally mandated safety regulations (Hamby 2020, Reynolds et al. 2018). One study of 19 miners identified patterns including lack of consistent ventilation maintenance in mines, miners wearing dust samplers incorrectly to purposely record a lower exposure amount, and miners leaving dust samplers in lower-dust areas of the mine (Reynolds et al. 2018). Disregarding or incorrectly following safety regulations designed to protect miners may increase the risk of CWP and contribute to higher rates of disease.

It is also worth noting that coal mine size in Appalachia can contribute to safety-related issues. Studies have suggested that smaller mines (with fewer than 50 employees) may have "limited knowledge of, and resources for, dust reduction and disease elimination," leading to higher prevalence of disease (Hendryx et al. 2013; Laney et al. 2012).

Appalachian coal mining communities. As discussed in Section 3.1.1, there is evidence that coal mining communities can suffer from higher rates of morbidity and mortality from diseases associated with coal dust exposure through coal mining, transportation, and processing (Huertas et al. 2014; Huertas et al. 2012; Mandal et al. 2012). This finding holds true when looking at Appalachia in particular. A study by Hendryx et al. (2008) found that areas of Appalachia with high levels of coal mining had significantly higher lung cancer mortality than the rest of Appalachia and the United States after adjusting for covariates such as smoking, poverty, and presence of current or former coal miners. Similarly, another study by Hendryx (2009) found significantly higher levels of heart disease, respiratory disease, and kidney disease among areas of Appalachia with high levels of mining compared to other populations. It should also be noted that covariates specific to the Appalachian region may impact prevalence of respiratory illnesses. These include behavioral, demographic, or cultural factors; tobacco use; and poverty (Hendryx 2009; Hendryx et al. 2008).

Residential coal use. While there is a lack of literature focusing on the health impacts of residential coal use in this region, the clear evidence of residential coal burning suggests this is a topic deserving further study. American Community Survey data from 2021 show almost 59,000 households across 13 Appalachian states use coal for home heating or cooking. A 2013 study estimated that up to 30% of homes in Central Appalachia use wood or coal as their primary fuel for heating (Paulin et al. 2013). Similarly, a 2014 study found concentrated solid fuel use in Appalachia (particularly parts of West Virginia and Kentucky) and the Four Corners region (New Mexico, Arizona, Utah, and Colorado), along with several other small pockets of concentration (Rogalsky et al. 2014). Additionally, rural areas with high proportions of solid fuel use in the home may "significantly contribute to ambient air pollution," increasing the risk of negative health impacts for the entire community (Rogalsky et al. 2014).

3.3.2. Context on the Navajo Nation

The available information on black lung disease in the Navajo Nation is limited, representing a gap in the available evidence.[24] This highlights the need for further research on black lung disease (and other respiratory illnesses associated with coal exposure) in the Navajo Nation. However, this gap in published literature is not evidence against the impact of black lung disease in this population. Li et al. (2018) noted that available studies "have consistently found an association between respiratory disease burdens and the use of wood and coal in several Navajo communities" (Li et al. 2018; Bunnell et al. 2010). Studies indicate high levels of residential coal use in the Navajo Nation and high public health burden among this population related to coal use and coal mining.

Coal mining. Coal mining has historically been important to the Navajo economy. Rekow (2019) describes the Navajo economy as "dependent on fossil fuel" with more than half of the nation's annual revenue coming from coal mining and thousands of Navajo Nation residents relying on coal mines for employment. As with coal miners in other parts of the country, this is associated with a higher incidence of black lung disease and other respiratory diseases among Navajo coal miners (Patel 2015).

Residential coal burning. It is well documented that many homes in the Navajo Nation use solid fuels such as wood and coal for cooking and heating (Li et al. 2018; Champion et al. 2017; Finkelman and Bunnell 2003; Finkelman and Simoni n.d.). Coal is commonly used in Navajo Nation homes due to its low cost and wide availability; in some cases, residents receive coal for free from nearby mines or can purchase coal inexpensively at local flea markets (Li et al. 2018; Champion et al. 2017; Bunnell et al. 2010; Bunnell and Garcia 2006; Finkelman and Simoni n.d.).

There are several factors that may exacerbate exposure to fine particulates from coal burning in Navajo Nation homes, such as residents closing controller dampers on coal stoves (to avoid heat loss), stoves in disrepair, and stoves that "were not designed to operate properly at the higher temperatures at which coal burns" (Bunnell et al. 2010). Bunnell et al. (2010) conducted $PM_{2.5}$ monitoring with 18 of 137 Navajo Nation households participating in a survey on home heating methods and noted observable cracks on coal stoves and evidence of soot on surfaces within the home, both signs of higher levels of particulate

[24] The team's search for relevant literature often turned up papers on uranium mining in the Navajo Nation. While the topic of uranium mining is out of scope for this project, it may be of potential interest for future research, as uranium miners face radon exposure that can lead to respiratory illnesses similar to black lung disease (Dr. Robert Finkelman, personal communication with authors, January 4, 2023).

matter 2.5 ($PM_{2.5}$) exposure. Based on the evidence that residential coal burning for cooking and heating creates indoor air pollution (Li et al. 2018; Champion et al. 2017; Bunnell et al. 2010; Finkelman and Bunnell 2003; Finkelman and Simoni n.d.) as discussed in Section 3.1.2, it is not unreasonable to suggest that the respiratory illnesses associated with respiratory coal dust exposure are likely present in the Navajo Nation.

Outdoor air pollution. In addition to indoor air pollution due to residential coal use, the town of Shiprock, New Mexico, in the Navajo Nation experiences "noticeable amounts of smog" due to its vicinity to several coal-fired power plants (Bunnell et al. 2010).[25] Because they were built before the legislation's effective date, these power plants are exempt from Environmental Protection Agency regulation under the 1990 American Clean Air Act (Bunnell et al. 2010). One study examining hospital records from April 1997 to December 2002 among Navajo residents seen at the Northern Navajo Medical Center Indian Health Services (IHS) Hospital founds that of all 37 communities in the IHS's Shiprock Service Area, residents in the town of Shiprock were at greater risk for respiratory disease than residents of Navajo Nation communities not subject to such smog (Bunnell et al. 2010).

Public health burden of coal in the Navajo Nation. Taken together, coal mining, residential coal burning, and coal-related outdoor air pollution lead to a high public health burden in the Navajo Nation. As Li et al. (2018) note, the evidence of higher rates of exposure to $PM_{2.5}$ due to residential coal use "may be major contributing factors to public health burdens observed in the Navajo Nation, such as the higher death rates due to cardiovascular and respiratory illness compared to the rest of the US." The literature shows that Navajo people suffer "high levels of respiratory disease," despite low rates of cigarette smoking (Bunnell and Garcia 2006). When compared with the overall U.S. population, the Navajo Nation and other Native Americans "suffer disproportionately from respiratory morbidity" (Finkelman and Simoni n.d.).

Additionally, it is possible that the public health burden of coal in the Navajo Nation is not fully documented, given underreporting due to lack of healthcare access and reluctance to trust healthcare providers. In the Navajo Nation, the Navajo Area Indian Health Service—the region's primary healthcare provider—has 0.91 hospital beds per 1,000 people, compared to an average of 2.76 beds per 1,000 people across the rest of the United States (Arambula Solomon et al. 2022; American Hospital Association 2022; Census Bureau 2022a). Looking at Native American populations more broadly, many Native Americans do not have health insurance or must travel long distances to reach the nearest healthcare provider (Arambula Solomon et al. 2022; Whitney 2017). Lack of trust may be another reason for the underreporting of health problems, as several studies have cited lower levels of trust in healthcare providers among Native Americans compared to other populations (Guadagnolo et al. 2009; Hunt et al. 2005).

3.3.3. Comparing prevalence across county types and populations of interest

This section presents the findings from a series of statistical hypothesis tests to compare different groups in the data; namely, county types and populations of interest. Specifically, the team sought to draw conclusive evidence that two types of counties differ significantly among black lung cases and

[25] While the San Juan Mine and the coal-powered San Juan Generating Station recently shut down in 2022 (Moses 2022; Robinson-Avila 2022), the APS Four Corners power plant continues to operate, albeit at a reduced rate (Randazzo 2021).

deaths. The following county types and populations of interest were tested for statistical differences in black lung cases and deaths, both overall and considering population size:

- Coal county status (coal versus non-coal);
- Historical coal status (current versus former coal counties);
- Geography (Appalachia versus the Navajo Nation versus other counties);
- Mine type (surface versus underground exposure); and
- Residential use (counties with residential use versus no residential use).

The team tested a hypothesis of equality, namely, the average prevalence is equal between the county types and populations of interest. Specifically, we performed a two-tailed *t*-test against the null hypothesis at 95% confidence that the average number of cases/deaths is equal between sub-populations. Each table below summarizes the results of these tests with a brief accompanying explanation.

Coal counties compared to non-coal counties. As shown in Table 12, the average coal county has 47.34 black lung cases overall, or 1.20 cases per 1,000 residents. Non-coal counties, on the other hand, have nearly zero black lung cases overall and fewer than 1 case per 100,000 residents. In terms of both counts and rates, coal counties have significantly more black lung cases than non-coal counties. The same finding is true about black lung deaths; there are roughly 38 deaths per 100,000 residents in coal counties compared to just 1 per 100,000 in non-coal counties. This does not necessarily mean that simply residing in a coal county leaves one susceptible to black lung disease, but rather that these counties include some key factors that may lead to these diseases, such as surrounding industries or frequent exposure to coal dust.

Research Question 2: How does black lung incidence compare between current, former, and non-coal mining communities?

Answer: Black lung disease is most prevalent among coal counties, specifically those that have maintained continuous production since the 1970s and 1980s. Those that have stopped mining coal have significantly lower black lung rates, and non-coal counties rarely report any black lung cases or deaths.

Table 12: Comparison of difference in black lung counts and rates, coal versus non-coal counties

	Coal	Non-Coal	*p*-Value	Statistically Significant?
Sample size	286	2,850	-	-
Avg. cases	47.34	0.02	0.00	Yes
Avg. cases per 1,000	1.20	< 0.01	0.00	Yes
Avg. deaths	16.82	2.09	0.00	Yes
Avg. deaths per 1,000	0.38	0.01	0.00	Yes

Sources: Center for Disease Control and Prevention (CDC) Enhanced Coal Workers' Health Surveillance Program (ECWHSP) 1970-2014, CDC Wide-ranging ONline Data for Epidemiologic Research (WONDER) database 1999-2020, and U.S. Census Bureau American Community Survey (ACS) 1-year estimates. Note: "Coal" and "Non-coal" county definitions used in the table are original and were created as part of this research.

Current coal counties compared to former coal counties. Of the 286 coal counties, current coal counties (N = 172) have significantly more black lung cases and deaths than former coal counties (N = 114), as shown in Table 13. As discussed in Section 3.1.3, black lung disease can take months or even years to develop. Therefore, we assume a lagged effect between exposure to coal dust and development of black lung disease. But even considering that lagged effect, current coal counties far outnumber former counties in terms of black lung cases and deaths. This finding is likely due to continued coal industry presence in many of these counties. Of the 172 current coal counties, almost all of them (154) were involved in the coal industry in the 1970s and 1980s as well. This means they mined coal in the past—like former coal counties—but also continue to do so today. This continued exposure to coal dust leaves more residents vulnerable to black lung disease than those where coal mining has ceased.

Table 13: Comparison of difference in black lung counts and rates, current versus former coal counties

	Current Coal	Former Coal	*p*-Value	Statistically Significant?
Sample size	172	114	-	-
Avg. cases	76.59	3.20	0.00	Yes
Avg. cases per 1,000	1.92	0.12	0.00	Yes
Avg. deaths	25.73	3.38	0.00	Yes
Avg. deaths per 1,000	0.60	0.04	0.00	Yes

Sources: Center for Disease Control and Prevention (CDC) Enhanced Coal Workers' Health Surveillance Program (ECWHSP) 1970-2014, CDC Wide-ranging ONline Data for Epidemiologic Research (WONDER) database 1999-2020, and U.S. Census Bureau American Community Survey (ACS) 1-year estimates. Note: "Current" and "Former" coal county definitions used in the table are original and were created as part of this research.

Appalachia, the Navajo Nation, and other counties. Of the 103 counties that report black lung cases between 1970-2014, Appalachia accounts for 79 (77%) of them, mostly from North and Central Appalachia (26 and 50, respectively). A similar result is true among black lung deaths, specifically deaths from CWP. As shown in Table 14, the team found that Appalachian counties have significantly higher black lung case (0.70 per 1,000) and death rates (0.24 per 1,000) than the rest of the United States.

Research Question 3: Are black lung incidences and deaths more prevalent in the Navajo Nation or Appalachia than other parts of the United States?

Answer: Appalachia has significantly more black lung cases and deaths than the rest of the United States, both nominally and controlling for population size. Results for the Navajo Nation are inconclusive due to a lack of available data in these counties.

Table 14: Comparison of difference in black lung counts and rates, Appalachia versus other counties

	Appalachia	Other	*p*-Value	Statistically Significant?
Sample size	422	2,708	-	-
Avg. cases	28.79	0.53	0.00	Yes
Avg. cases per 1,000	0.70	0.02	0.00	Yes
Avg. deaths	10.88	2.28	0.00	Yes
Avg. deaths per 1,000	0.24	0.01	0.00	Yes

Sources: Center for Disease Control and Prevention (CDC) Enhanced Coal Workers' Health Surveillance Program (ECWHSP) 1970-2014, CDC Wide-ranging ONline Data for Epidemiologic Research (WONDER) database 1999-2020, and U.S. Census Bureau American Community Survey (ACS) 1-year estimates. Note: Navajo Nation counties are excluded from "other counties."

As previously noted, ECWHSP does not report county-level results with less than 10 confirmed cases, for privacy purposes. As shown in Table 15, none of the Navajo Nation's six counties report more than 10 black lung cases.[26] Furthermore, only one county (San Juan, New Mexico) reported black lung deaths between 1999 and 2020. This lack of evidence makes it impossible to statistically conclude whether the Navajo Nation is different from either Appalachia or the rest of the United States. The research team attempted to estimate the amount of black lung not reported within the Navajo Nation. This exercise can be found in Appendix C.

[26] This may be a sign of data suppression in the Navajo Nation, which is discussed in the limitations in Section 3.4.

Table 15: Comparison of difference in black lung counts and rates, Navajo Nation versus other counties

	Navajo Nation	Other	*p*-Value	Statistically Significant?
Sample size	6	2,708	-	-
Avg. cases	0.00	0.53	0.87	No
Avg. cases per 1,000	0.00	0.02	0.86	No
Avg. deaths	1.67	2.28	0.86	No
Avg. deaths per 1,000	0.01	0.01	0.97	No

Sources: Center for Disease Control and Prevention (CDC) Enhanced Coal Workers' Health Surveillance Program (ECWHSP) 1970-2014, CDC Wide-ranging ONline Data for Epidemiologic Research (WONDER) database 1999-2020, and U.S. Census Bureau American Community Survey (ACS) 1-year estimates. Note: Appalachian counties are excluded from "other counties."

Surface mines compared to underground mines. As shown in Table 16, black lung cases are significantly more prevalent in underground coal mines (11 cases per 100,000) than surface coal mines (less than 1 case per 100,000). Note that this finding alone does not mean that underground mines cause black lung disease and surface mines do not. It also may be reasonable to expect higher black lung rates among underground mines because these likely have more concentrated exposure to coal dust as opposed to open-air working conditions in surface mines.[27]

Table 16: Comparison of difference in black lung counts and rates, surface versus underground mines

	Surface	Underground	*p*-Value	Statistically Significant?
Sample size	3,136	3,136	-	-
Avg. cases	0.09	4.24	0.00	Yes
Avg. cases per 1,000	< 0.01	0.11	0.00	Yes

Sources: Center for Disease Control and Prevention (CDC) Enhanced Coal Workers' Health Surveillance Program (ECWHSP) 1970-2014, CDC Wide-ranging ONline Data for Epidemiologic Research (WONDER) database 1999-2020, Mine Safety and Health Administration (MSHA) Mines Dataset, and U.S. Census Bureau American Community Survey (ACS) 1-year estimates.

Residential coal use. As shown in Table 17, counties whose households use coal have significantly higher black lung case and death rates than those that do not. There are about 0.26 black lung cases per 1,000 residents (26 per 100,000) compared to just 0.03 cases per 1,000 (3 cases per 100,000) among counties that do not use coal in the home. Similarly, counties with residential coal use report 0.10 deaths per 1,000 residents (10 deaths per 100,000) versus 0.01 deaths (1 per 100,000) in counties without residential coal use. In both cases, these results are statistically significant. Note that this correlation and statistical difference do not imply causation between residential use and black lung disease, only that these two groups are conclusively different from one another.[28]

Research Question 4: How does residential coal burning correlate with black lung cases and deaths?

Answer: There is a correlation between residential coal use and black lung cases/deaths. Further, there are significantly more cases and deaths from black lung in residential coal burning areas than other parts of the United States.

[27] Black lung death data are not tracked according to the type of mine in which the miner worked.
[28] For additional information on correlations between residential coal use and other coal metrics, see Appendix B.

Table 17: Comparison of difference in black lung counts and rates, residential coal use

	Yes	No	*p*-Value	Statistically Significant?
Sample size	1,105	2,031	-	-
Avg. cases	10.61	0.92	0.00	Yes
Avg. cases per 1,000	0.26	0.03	0.00	Yes
Avg. deaths	8.34	0.77	0.00	Yes
Avg. deaths per 1,000	0.10	0.01	0.00	Yes

Sources: Center for Disease Control and Prevention (CDC) Enhanced Coal Workers' Health Surveillance Program (ECWHSP) 1970-2014, CDC Wide-ranging ONline Data for Epidemiologic Research (WONDER) database 1999-2020, and U.S. Census Bureau American Community Survey (ACS) 1-year estimates. Note: Residential coal use definitions for counties used in the table are original and were created as part of this research.

3.4. Limitations

This section describes five key limitations which should be considered when reviewing the findings of this study.

Diagnosing black lung*:* The term "black lung" is somewhat ambiguous as it is not explicitly tied to a medical diagnosis code in the International Classification of Diseases (ICD) system. Therefore, black lung cases and deaths are measured according to the definitions used by each federal program collecting its own data. Black lung cases are recorded by CDC's Enhanced Coal Workers' Health Surveillance Program (ECWHSP) according to the ICD-10 diagnosis code J60, coal workers' pneumoconiosis (CWP). Black lung deaths, on the other hand, are measured according to a more expansive list of ICD-10 codes which includes CWP, silicosis, asbestosis, Berylliosis, and other unspecified pneumoconiosis (J60, J61, J62, J62.8, J63.2, J64). Because of these distinct definitions, metrics for black lung cases and deaths are reported separate from one another. Section 3.1.3 also discusses this difference in collection in more detail.

Measuring black lung incidence: There are several programs that aim to measure black lung prevalence across the United States, both federally and privately funded. Diagnosis of respiratory disease due to coal mine dust exposure requires respiratory symptoms, medical tests such as lung imaging and pulmonary function testing, and a detailed history of exposure. Because the symptoms and respiratory illnesses associated with exposure to coal dust can easily resemble other respiratory illnesses, the only way to conclusively pinpoint the cause to coal dust is through a well-documented history of exposure or through autopsies.

Time frames: Black lung case and death statistics were collected by different entities and across different time periods. This makes it difficult to directly compare living cases to deaths. Therefore, each metric is analyzed separately. Furthermore, when black lung cases and deaths are reported per 1000 residents, the result is based upon the population in the final year of the data collection period. Black lung cases per 1000 residents are reported according to 2014 population estimates; black lung deaths per 1000 residents are reported according to 2020 population.

Meanwhile, data from Census, CDC, EIA, and MSHA are all collected over different periods with irregular periods of collection. The team could not construct a panel dataset for this analysis without exceeding the time and budget constraints of the project. Therefore, the team used discrete points in time (i.e., years) to approximate past/present time periods in the data. This is especially relevant in the classification of “current” and “former” coal counties.

Data suppression: Black lung case and death data are collected by ECWHSP and WONDER, two publicly funded projects. They do not list data for districts, states, or counties with less than 10 examined miners to protect individual privacy. This means that counties with nine or fewer black lung cases, for example, show up in the data as zero cases, which is a form of data suppression. This practice makes it impossible to identify counties with less than 10 black lung cases, making the data more difficult to interpret. Note that suppression applies to the number of *total* cases across the county regardless of disease severity. For instance, some counties report 10 total black lung cases, nine of which are not severe, one which is very severe. Because there are 10 total cases across the county, the level of severity will not be suppressed, meaning those distinct measurements can be disentangled within the county.

Data underreporting: In addition to data suppression, research suggests that counties defined as the Navajo Nation likely have a number of underreported black lung cases. Members of the Navajo Nation may have undiagnosed cases of black lung or choose not to participate in federally sponsored programs. While this underreporting is difficult to quantify, the researchers believe that there are more black lung cases and deaths than appear in the data. To help estimate this potential undercounting, the research team used a series of regression models to estimate the number of black lung cases and deaths in each county. The results of those models can be found in Appendix C. Finally, the research team only leveraged publicly available data for this study. No private sources or datasets were procured. Therefore, there is also a risk of data underreporting for black lung disease for coal workers and their families who did not have data submitted to ECWHSP or CDC’s WONDER database.

4. Conclusions and next steps

This section highlights the key takeaways informing each study research question and identifies some potential next steps for future research.

4.1. Key takeaways by research question

The key takeaways from quantitative analysis and the literature are presented in Table 18 and organized by research question.

Table 18: Key takeaways by research question

RQ1: What is the rate of black lung disease across the United States?	Report Section
• On average, there are 4.34 cumulative cases (1970-2014) and 3.44 cumulative deaths (1999-2020) attributable to black lung disease per county across the United States. These results are counts and therefore not subject to population size.	3.2.1
• The prevalence of black lung disease is highly concentrated in specific areas of the country, such as Appalachia, where the statistics rise to 28.79 cumulative cases and 10.88 cumulative deaths, on average. 103 counties out of 3,136 report black lung cases during the entire collection period (1970-2014) while 333 counties report black lung deaths (1999-2020). These results are counts and not subject to population size.	3.3.3
• Coal workers' pneumoconiosis (CWP) prevalence has been increasing in the United States since the 1990s.	3.1.1
• Undercounts in the caseload of black lung disease in the United States may be attributed to factors including low uptake of Coal Workers' Health Surveillance Program screenings and low compensation approval rates.	3.1.3
RQ1.1: What is the total number and rate per 1,000 residents of black lung cases?	
• There are 0.11 black lung cases per 1,000 U.S. residents, or 11 cases per 100,000.*	3.2.1
RQ1.2: What is the total number and rate per 1,000 residents of black lung deaths?	
• There are 0.04 black lung deaths per 1,000 U.S. residents, or four cases per 100,000.*	3.2.2
RQ2: How does black lung incidence compare between current, former, and non-coal mining communities?	
• Black lung disease is most prevalent among coal counties,^ especially those that have maintained coal mining practices since the 1970s and 1980s.	3.3.3
• Former coal counties have significantly lower black lung rates than current coal counties.	3.3.3
• Non-coal counties rarely report any cases or deaths attributable to black lung disease. Deaths in these counties most often come from asbestosis, which may risk overcounting.	3.3.3
RQ3: Are black lung cases and deaths more prevalent in the Navajo Nation or Appalachia than other parts of the United States?	
• Appalachia has significantly more black lung cases and deaths than the rest of the United States, both overall and controlling for population size.	3.3.3
• Zero cases and just 10 deaths were reported in the Navajo Nation during the respective collection periods (1970-2014 and 1999-2020). Therefore, hypothesis testing between the Navajo Nation and other regions proved inconclusive. These results may be a function of underreporting in the Navajo Nation.	3.3.3
RQ4: How does residential coal burning correlate with black lung cases and deaths?	
• There is a correlation between residential coal use and black lung cases and deaths. However, this correlation does not imply a causal relationship between residential use and black lung disease.	3.3.3
• There are significantly more cases and deaths from black lung in residential coal burning areas than other parts of the United States.	3.3.3

RQ4: How does residential coal burning correlate with black lung cases and deaths?	Report Section
• Residential coal burning is likely a confounding factor with other black lung characteristics, such as the number of local underground mines or family members working in the coal industry.	3.3.3

* Results are cumulative across the respective data collection period (i.e., black lung cases 1970-2014, black lung deaths 1999-2020). Population estimates are based on 2014 and 2020 data, respectively.
^ Summit classified a county as a coal county if any of the following four criteria are met: county has 1+ mines in either 1983 or 2020, county produced 1+ short tons of coal in 1983 or 2020, there is 1+ coal miners residing in the county in 1986 or 2020, or workers in the county are exposed to an average of 1+ hours to coal mine dust each week between 1970 and 2020.
Note: Key findings were derived based on literature review sources (see Appendix A) as well as publicly available data sources for quantitative analyses (see Section 2.2.1).

4.2. Potential next steps

The results of this study show that black lung disease due to unsafe practices in coal mining, residential coal burning, and air pollution through coal processing and transportation remains at concerning levels in the United States, as highlighted by the rising rates of CWP since the 1990s. In addition, Appalachia has significantly more black lung cases and deaths than the rest of the United States, both in terms of counts and controlling for population size. This study has identified several areas of potential future research to help MSHA understand the impact of coal on health, both in terms of specific populations of interest and types of negative health impacts.

The Navajo Nation. The literature review findings suggest that the public health burden of coal in the Navajo Nation is high, due to the historic importance of coal mining to the local economy and high levels of residential coal use. However, lack of data in the Navajo Nation counties led to inconclusive analytic results. Given evidence that this population may experience a disproportionate negative impact from coal compared to the rest of the country, MSHA could consider future research to address the lack of public health data available for analysis. Specific next steps to design and launch such a study will be detailed in a separate memo on the Navajo Nation–specific findings from this study.

Silicosis. MSHA recently announced a proposed rule change related to 30 CFR § 56, 57, 70, 71, 72, and 90 which will address health hazards from silica dust exposure, which can lead to severe illness including silicosis and PMF (MSHA 2023). This rule change was announced after the literature review and statistical analysis for this report had been completed. As described in Section 3.1.4, the team opted to use the list of ICD codes in Table 8 to capture a more expansive view of black lung disease than the strict CWP tracking done by ECWHSP. While this included ICD-10 code J62 (pneumoconiosis due to dust containing silica, also known as silicosis), none of the analysis presented in this report is specifically focused on silicosis. Similarly, sources that solely discussed silicosis were not reviewed as part of the literature review. Given the announced rule change, MSHA may consider conducting a follow-on study with similar research questions that focus specifically on silicosis—rather than on black lung disease more broadly—to supplement the findings of this report. Such a study would require a dataset scan like the one conducted for this study, to determine what data is publicly available on silicosis diagnoses.

APPENDIX A Sources cited

American Hospital Association. 2022. "Fast Facts on U.S. Hospitals, 2022." AHA.org (website). Last modified January 2022. https://www.aha.org/statistics/fast-facts-us-hospitals.

Arambula Solomon, Teshia G., Rachel Rose Bobelu Starks, Agnes Attakai, Fatima Molina, Felina Cordova-Marks, Michelle Kahn-John, Chester L. Antone, Miguel Flores, and Francisco Garcia. 2022. "The Generational Impact Of Racism On Health: Voices From American Indian Communities." *Health Affairs* 41, no. 2 (February). https://doi.org/10.1377/hlthaff.2021.01419.

Arif, Ahmed A., Rajib Paul, Eric Delmelle, Claudio Owusu, and Oluwaseun Adeyemi. 2020. "Estimating the prevalence and spatial clusters of coal workers' pneumoconiosis cases using medicare claims data, 2011-2014." *American Journal of Industrial Medicine* 63, no. 6 (March 9, 2020): 478–83. https://doi.org/10.1002/ajim.23104.

Arnold, Carrie. 2016. "A Scourge Returns: Black Lung in Appalachia." *Environmental Health Perspectives* 124, no. 1 (January 1, 2016): A13–18. https://doi.org/10.1289/ehp.124-A13.

Barone-Adesi, Francesco, Robert S. Chapman, Debra T. Silverman, Xinghzhou He, Wei Hu, Roel Vermeulen, Bofu Ning, Joseph F. Fraumeni Jr., Nathaniel Rothman, and Qing Lan. 2012. "Risk of lung cancer associated with domestic use of coal in Xuanwei, China: retrospective cohort study." *BMJ* 345 (August 30, 2012): e5414. https://doi.org/10.1136/bmj.e5414.

Blackley, David J., Cara N. Halldin, and A. Scott Laney. 2018. "Continued Increase in Prevalence of Coal Workers' Pneumoconiosis in the United States, 1970–2017." *American Journal of Public Health* 108, no. 9 (September): 1220–22. https://doi.org/10.2105/AJPH.2018.304517.

Braithwaite, Isobel, Shuo Zhang, James B. Kirkbride, David P. J. Osborn, and Joseph F. Hayes. 2019. "Air Pollution (Particulate Matter) Exposure and Associations with Depression, Anxiety, Bipolar, Psychosis and Suicide Risk: A Systematic Review and Meta-Analysis." *Environmental Health Perspectives* 127, no. 12 (December 18, 2019): 126002. https://doi.org/10.1289/EHP4595.

Buchanan, Susan, Erica Burt, and Peter Orris. 2014. "Beyond black lung: scientific evidence of health effects from coal use in electricity generation." *Journal of Public Health Policy* 35 (2014): 266–77. https://doi.org/10.1057/jphp.2014.16.

Bunnell, Joseph E., and Linda V. Garcia. 2006. "Navajo coal and air quality in Shiprock, New Mexico." USGS Numbered Series: 2006-2094. USGS. https://doi.org/10.3133/fs20063094.

Bunnell, Joseph E., Linda V. Garcia, Jill M. Furst, Harry Lerch, Ricardo A. Olea, Stephen E. Suitt, and Allan Kolker. 2010. "Navajo Coal Combustion and Respiratory Health Near Shiprock, New Mexico." *Journal of Environmental and Public Health* 2010 (June 30, 2010): e260525. https://doi.org/10.1155/2010/260525.

Cartwright, Elizabeth. 2016. "Mining and Its Health Consequences: From Matewan to Fracking." *A Companion to Anthropology and Environmental Health* Chapter 20 (April 15, 2016): 417–34. https://doi.org/10.1002/9781118786949.ch20.

Cecil, Mollie M. 2021. "Doctors, Miners, and Black Lung: A transatlantic comparison of organized medicine's role in the fight for black lung recognition in West Virginia and Wales." MA, West Virginia University Libraries. *Graduate Theses, Dissertations, and Problem Reports* 8093. https://doi.org/10.33915/etd.8093.

Centers for Disease Control and Prevention. 2023. "CDC WONDER." CDC Wide-ranging ONline Data for Epidemiologic Research (WONDER). https://wonder.cdc.gov.

Center for Disease Control (CDC) National Institute for Occupational Safety & Health (NIOSH). 2020. "Enhanced Coal Workers' Health Surveillance Program (ECWHSP) Data Query System." Enhanced Coal Workers' Health Surveillance Program. https://webappa.cdc.gov/ords/cwhsp-database.html#queryForm.

Champion, W. M., P. H. Charley, B. Klein, K. Stewart, P. A. Solomon, and L. D. Montoya. 2017. "Perception, culture, and science: a framework to identify in-home heating options to improve indoor air quality in the Navajo Nation." *Science of The Total Environment* 580 (February 15, 2017): 297–306. https://doi.org/10.1016/j.scitotenv.2016.11.053.

CMS (U.S. Centers for Medicare & Medicaid Services). 2021. "2016 ICD-10-CM and GEMs." CMS.gov (website). Last modified December 1, 2021. https://www.cms.gov/Medicare/Coding/ICD10/2016-ICD-10-CM-and-GEMs.

CMS. 2022. "ICD Code Lists." CMS.gov (website). Last modified September 28, 2022. https://www.cms.gov/medicare/coordination-benefits-recovery-overview/icd-code-lists.

Cohn, Corey A., Richard Laffers, Sanford R. Simon, Thomas O'Riordan, and Martin AA Schoonen. 2006. "Role of pyrite in formation of hydroxyl radicals in coal: possible implications for human health." *Particle and Fibre Toxicology* 3, no. 1 (December 19, 2006): 16. https://doi.org/10.1186/1743-8977-3-16.

Cortes-Ramirez, Javier, Suchithra Naish, Peter D Sly, and Paul Jagals. 2018. "Mortality and Morbidity in Populations in the Vicinity of Coal Mining: A Systematic Review." *BMC Public Health* 18, no. 721 (June 11, 2018). https://doi.org/10.1186/s12889-018-5505-7.

DOL. 2023. "Black Lung Program Statistics." Black Lung Program Statistics. Accessed January 24, 2023. https://www.dol.gov/agencies/owcp/dcmwc/statistics/PartCClaimsDecisions.

Finkelman, Robert B., and Fiorella V. Simoni. n.d. "Relationship of Indoor and Ambient Air Quality to Respiratory Diseases in the Navajo Nation." Unpublished.

Finkelman, Robert B., Amy Wolfe, and Michael S. Hendryx. 2021. "The future environmental and health impacts of coal." *Energy Geoscience* 2, no. 2 (April 1, 2021): 99–112. https://doi.org/10.1016/j.engeos.2020.11.001.

Finkelman, Robert, and Joseph E. Bunnell. 2003. "Integrated Approaches to Understanding the Relationship of Coal to Human and Ecosystem Health." *Program and Abstracts for the Twentieth Annual Meeting of The Society for Organic Petrology*, Open-File Report, 20 (September): 61–64.

Finkelman, Robert, William Orem, Vincent Castranova, Calin A. Tatu, Harvey E. Belkin, Baoshan Zheng, Harry Lerch, Susan V Maharaj, and Anne L. Bates. 2002. "Health Impacts of Coal and Coal Use: Possible Solutions." *International Journal of Coal Geology* 50: 425–43.

Galeone, Carlotta, Claudio Pelucchi, Carlo La Vecchia, Eva Negri, Cristina Bosetti, and Jinfu Hu. 2008. "Indoor Air Pollution from Solid Fuel Use, Chronic Lung Diseases and Lung Cancer in Harbin, Northeast China." *European Journal of Cancer Prevention* 17, no. 5: 473–78.

Guadagnolo, B. Ashleigh, Kristin Cina, Petra Helbig, Kevin Molloy, Mary Reiner, E. Francis Cook, and Daniel G. Petereit. 2009. "Medical mistrust and less satisfaction with health care among Native Americans presenting for cancer treatment." *Journal of Healthcare for the Poor and Underserved* 20, no. 1 (February): 210–26. 10.1353/hpu.0.0108.

Hamby, Chris. 2020. *Soul Full of Coal Dust: A Fight for Breath and Justice in Appalachia*. United Kingdom: Hachette Book Group.

Hamby, Chris, Brian Ross, and Matthew Mosk. 2013. "Johns Hopkins medical unit rarely finds black lung, helping coal industry defeat miners' claims." The Center for Public Integrity (October 30, 2013). https://publicintegrity.org/environment/johns-hopkins-medical-unit-rarely-finds-black-lung-helping-coal-industry-defeat-miners-claims/.

Harrington, Andrea D., Shavonne Hylton, and Martin A. A. Schoonen. 2012. "Pyrite-driven reactive oxygen species formation in simulated lung fluid: implications for coal workers' pneumoconiosis." *Environmental Geochemistry and Health* 34, no. 4 (August 1, 2012): 527–38. https://doi.org/10.1007/s10653-011-9438-7.

Hendryx, Michael. 2009. "Mortality from heart, respiratory, and kidney disease in coal mining areas of Appalachia." *International Archives of Occupational and Environmental Health* 82, no. 2 (January 2009): 243–49. https://doi.org/10.1007/s00420-008-0328-y.

Hendryx, Michael, Kathryn O'Donnell, and Kimberly Horn. 2008. "Lung cancer mortality is elevated in coal-mining areas of Appalachia." *Lung Cancer* 62, no. 1 (October 2008): 1–7. https://doi.org/10.1016/j.lungcan.2008.02.004.

Hendryx, Michael, Cynthia Persily, Jamison Conley, and Evan Fedorko. 2013. "Rates of Black Lung Disease in Relationship to Black Lung Treatment Centers." West Virginia Rural Health Research Center, 2013. https://www.ruralhealthresearch.org/mirror/5/521/black-lung-clinics-final-report-final-073013.pdf.

Hosgood, H. Dean, Paolo Boffetta, Sander Greenland, Yuan-Chin Amy Lee, John McLaughlin, Adeline Seow, Eric J. Duell, et al. 2010. "In-Home Coal and Wood Use and Lung Cancer Risk: A Pooled Analysis of the International Lung Cancer Consortium." *Environmental Health Perspectives* 118, no. 12 (December 1, 2010): 1743–47. https://doi.org/10.1289/ehp.1002217.

Hosgood, H Dean, III, Hu Wei, Amir Sapkota, Imran Choudhury, Nigel Bruce, Kirk R Smith, Nathaniel Rothman, and Qing Lan. 2011. "Household coal use and lung cancer: systematic review and meta-analysis of case–control studies, with an emphasis on geographic variation." *International Journal of Epidemiology* 40, no. 3 (June 2011): 719–28. https://doi.org/10.1093/ije/dyq259.

Huang, Xi, Terry Gordon, William N. Rom, and Robert B. Finkelman. 2006. "Interaction of Iron and Calcium Minerals in Coals and their Roles in Coal Dust-Induced Health and Environmental Problems." *Reviews in Mineralogy and Geochemistry* 64, no. 1 (January 1, 2006): 153–78. https://doi.org/10.2138/rmg.2006.64.6.

Huertas, J. I., M. E. Huertas, G. Cervantes, and J. Díaz. 2014. "Assessment of the natural sources of particulate matter on the opencast mines air quality." *Science of The Total Environment* 493 (September 15, 2014): 1047–55. https://doi.org/10.1016/j.scitotenv.2014.05.111.

Huertas, José I., María E. Huertas, Sebastián Izquierdo, and Enrique D. González. 2012. "Air quality impact assessment of multiple open pit coal mines in northern Colombia." *Journal of Environmental Management* 93, no. 1 (January 1, 2012): 121–29. https://doi.org/10.1016/j.jenvman.2011.08.007.

Hunt, Kelly A., Ayorkor Gaba, and Risa Lavizzo-Mourey. 2005. "Racial and Ethnic Disparities and Perceptions of Health Care: Does Health Plan Type Matter?" *Health Services Research* 40, no. 2 (April 2005): 551–576. 10.1111/j.1475-6773.2005.00372.

ILO (International Labour Office of Geneva). 2011. *Guidelines for the Use of the ILO International Classification of Radiographs of Pneumoconioses*. Revised Edition. Occupational Safety and Health Series 22. International Labour Organization.

Karatela, Shamshad, Samantha Caruana, and Gunther Paul. 2022. "Prevalence of Respiratory Disease in the Population of Queensland Communities in Proximity to Coal Mines and Coal Mining Activities." *International Journal of Community Medicine and Public Health* 9, no. 7 (July 1, 2022): 3014–22. https://doi.org/10.18203/2394-6040.ijcmph20221776.

Kerimray, Aiymgul, Luis Rojas-Solórzano, Mehdi Amouei Torkmahalleh, Philip K. Hopke, and Brian P. Ó Gallachóir. 2017. "Coal use for residential heating: patterns, health implications and lessons learned." *Energy for Sustainable Development* 40 (October 2017): 19–30. https://doi.org/10.1016/j.esd.2017.05.005.

Kuempel, E. D., M. D. Attfield, V. Vallyathan, N. L. Lapp, J. M. Hale, R. J. Smith, and V. Castranova. 2003. "Pulmonary inflammation and crystalline silica in respirable coal mine dust: Dose response." *Journal of Biosciences* 28, no. 1 (February 1, 2003): 61–69. https://doi.org/10.1007/BF02970133.

Kurth, Laura, and Megan Casey. 2020. "Pneumoconiosis ICD-CM Diagnosis Codes on Medicare Claims for Federal Black Lung Program Beneficiaries." *Annals of the American Thoracic Society* 17, no. 7 (July 2020): 904–6. https://doi.org/10.1513/AnnalsATS.202001-037RL.

Laney, A. Scott, and Michael D. Attfield. 2010. "Coal workers' pneumoconiosis and progressive massive fibrosis are increasingly more prevalent among workers in small underground coal mines in the United States." *Occupational and Environmental Medicine* 67, no. 6 (June 2010): 428–31. https://doi.org/10.1136/oem.2009.050757.

Laney, A. Scott, Edward L. Petsonk, Janet M. Hale, Anita L. Wolfe, and Michael D. Attfield. 2012. "Potential Determinants of Coal Workers' Pneumoconiosis, Advanced Pneumoconiosis, and Progressive Massive Fibrosis Among Underground Coal Miners in the United States, 2005–

2009." *American Journal of Public Health* 102, no. S2 (May 2012): S279–83. https://doi.org/10.2105/AJPH.2011.300427.

Laney, A. Scott, and Michael D. Attfield. 2014. "Examination of Potential Sources of Bias in the US Coal Workers' Health Surveillance Program." *American Journal of Public Health* 104, no. 1 (January 2014): 165–70. https://doi.org/10.2105/AJPH.2012.301051.

Laney, A. Scott, and David N. Weissman. 2014. "Respiratory Diseases Caused by Coal Mine Dust." *Journal of Occupational and Environmental Medicine* 56 (October 2014): S18–22. https://doi.org/10.1097/JOM.0000000000000260.

Large, David J., Shona Kelly, Baruch Spiro, Linwei Tian, Longyi Shao, Robert Finkelman, Mingquan Zhang, et al. 2009. "Silica–Volatile Interaction and the Geological Cause of the Xuan Wei Lung Cancer Epidemic." *Environmental Science & Technology* 43, no. 23 (November 3, 2009): 9016–21. https://doi.org/10.1021/es902033j.

Li, Ning, Wyatt M. Champion, Jemal Imam, Damansher Sidhu, Joseph R. Salazar, Brian J. Majestic, and Lupita D. Montoya. 2018. "Evaluation of cellular effects of fine particulate matter from combustion of solid fuels used for indoor heating on the Navajo Nation using a stratified oxidative stress response model." *Atmospheric Environment* 182 (June 1, 2018): 87–96. https://doi.org/10.1016/j.atmosenv.2018.03.031.

Lockwood, Alan H. 2012. *The Silent Epidemic: Coal and the Hidden Threat to Health*. Massachusetts: MIT Press.

Mandal, K., A. Kumar, N. Tripathi, R. S. Singh, S. K. Chaulya, P. K. Mishra, L. K. Bandyopadhyay, and Council of Scientific and Industrial Research. 2012. "Characterization of different road dusts in opencast coal mining areas of India." *Environmental Monitoring and Assessment* 184, no. 6 (June 1, 2012): 3427–41. https://doi.org/10.1007/s10661-011-2197-1.

McCunney, Robert J., Peter Morfeld, and Stephen Payne. 2009. "What Component of Coal Causes Coal Workers' Pneumoconiosis?" *Journal of Occupational and Environmental Medicine* 51, no. 4 (April): 462–71. https://doi.org/10.1097/JOM.0b013e3181a01ada.

Moses, John R. 2022. "San Juan Mine sealed as Westmoreland lays off 60, moves to reclamation phase." *Farmington Daily Times,* December 1, 2022. https://www.daily-times.com/story/news/local/four-corners/2022/12/01/san-juan-mine-sealed-westmoreland-announces-60-more-layoffs/69690226007/.

MSHA. 2023. "US Department of Labor announces proposed rule to reduce silica dust exposure, better protect miners' health." DOL MSHA (website). Accessed August 1, 2023.

MSHA. n.d. "Accident Injuries Data Set." DOL MSHA (website). Accessed April 6, 2023. https://arlweb.msha.gov/OpenGovernmentData/OGIMSHA.asp.

MSHA. n.d. "End Black Lung – Act Now!" DOL MSHA (website). Accessed January 24, 2023. https://arlweb.msha.gov/S&HINFO/BlackLung/Homepage2009.asp#jump.

MSHA. n.d. "Mission." DOL MSHA (website). Accessed April 6, 2023. https://www.msha.gov/about/mission.

Navajo Nation. n.d. "History." The Official Site of the Navajo Nation (website). Accessed September 29, 2023. https://www.navajo-nsn.gov/History

NIOSH. 2020. "Study Syllabus for Classification of Radiographs of Pneumoconioses." CDC.gov (website). August 5, 2020. https://www.cdc.gov/niosh/learning/b-reader/clinical/lung/2.html.

Patel, Sumera. 2015. "Cheap Energy." *ESSAI* 13, no. 1: 32. https://dc.cod.edu/cgi/viewcontent.cgi?article=1547&context=essai.

Patra, Aditya Kumar, Sneha Gautam, and Prashant Kumar. 2016. "Emissions and human health impact of particulate matter from surface mining operation—A review." *Environmental Technology & Innovation* 5 (April 1, 2016): 233–49. https://doi.org/10.1016/j.eti.2016.04.002.

Paul, Rajib, Oluwaseun Adeyemi, and Ahmed A. Arif. 2022. "Estimating mortality from coal workers' pneumoconiosis among Medicare Beneficiaries with pneumoconiosis using binary regressions for spatially sparse data." *American Journal of Industrial Medicine* 65, no. 4 (February 8, 2022): 262–67. https://doi.org/10.1002/ajim.23330.

PBS NewsHour. 2006. "Rate of black lung disease among miners may be 10 times higher than reported." Video of news broadcast, December 16, 2016. https://www.youtube.com/watch?v=d-lasggfisg.

Perret, Jennifer L., Brian Plush, Philippe Lachapelle, Timothy S. C. Hinks, Clare Walter, Philip Clarke, Louis Irving, Pat Brady, Shyamali C. Dharmage, and Alastair Stewart. 2017. "Coal mine dust lung disease in the modern era." *Respirology* 22, no. 4 (March 30, 2017): 662–70. https://doi.org/10.1111/resp.13034.

Petsonk, Edward L., Cecile Rose, and Robert Cohen. 2013. "Coal Mine Dust Lung Disease. New Lessons from an Old Exposure." *American Journal of Respiratory and Critical Care Medicine* 187, no. 11 (June 1, 2013): 1178–85. https://doi.org/10.1164/rccm.201301-0042CI.

Potera, Carol. 2019. "Black Lung Disease Resurges in Appalachian Coal Miners." *AJN, American Journal of Nursing* 119, no. 4 (April): 14. https://doi.org/10.1097/01.NAJ.0000554534.63115.c4.

Ranavaya II, Mohammed I, Mohammed Ranavaya, and Natavoot Chongswatdi. 2020. "Black Lung: Old Disease, New Lessons." *Marshall Journal of Medicine* 6, no. 3 (July 31, 2020): 27. https://doi.org/10.33470/2379-9536.1290.

Randazzo, Ryan. 2021. "Coal-burning Four Corners Power Plant will cut back on operations, APS says." *AZCentral,* March 12, 2021. https://www.azcentral.com/story/money/business/energy/2021/03/12/aps-four-corners-power-plant-reduce-operations-one-generator/4655198001/.

Rekow, Lea. 2019. "Mapping a Bleak Future: The Mining Legacy of Navajo Nation." *On Active Grounds: Agency and Time in the Environmental Humanities*, 1–19. Wilfrid Laurier University Press.

Reynolds, Laura E., David J. Blackley, Jay F. Colinet, J. Drew Potts, Eileen Storey, Connie Short, Ron Carson, Kathleen A. Clark, A. Scott Laney, and Cara N. Halldin. 2018. "Work Practices and Respiratory Health Status of Appalachian Coal Miners With Progressive Massive Fibrosis." *Journal of Occupational and Environmental Medicine* 60, no. 11 (November): e575–81. https://doi.org/10.1097/JOM.0000000000001443.

Robinson-Avila, Kevin. 2022. "San Juan Generating Station to close this week." *Albuquerque Journa*l, September 26, 2022. https://www.abqjournal.com/2535435/san-juan-coal-plant-to-close-this-week.html.

Rogalsky, Derek K., Pauline Mendola, Tricia A. Metts, and William J. Martin II. 2014. "Estimating the number of low-income Americans exposed to household air pollution from burning solid fuels." *Environmental Health Perspectives* 122, no. 8: 806–10. http://dx.doi.org/10.1289/ehp.1306709.

Ross, M. H., and J. Murray. 2004. "Occupational respiratory disease in mining." *Occupational Medicine* 54, no. 5 (August 1, 2004): 304–10. https://doi.org/10.1093/occmed/kqh073.

Royal, Kenneth. 2019. "Prevalence rates for black lung epidemic likely underestimated for coal miners in Appalachia." *Social Determinants of Health* 5, no. 1 (August 15, 2019): 62–4. https://doi.org/10.22037/sdh.v5i1.25717.

Shriver, Thomas E., and Aysha Bodenhamer. 2018. "The enduring legacy of black lung: environmental health and contested illness in Appalachia." *Sociology of Health & Illness* 40, no. 8 (June 28, 2018): 1361–75. https://doi.org/10.1111/1467-9566.12777.

Sisk, Taylor. 2023. "Black lung resurgence drives push to protect coal miners against silica dust." Kaiser Health News, March 14, 2023. https://khn.org/news/article/coal-miners-silica-dust-regulations-black-lung-resurgence/.

Somers, Sarah. 2017. "Congressional Plans to Repeal the ACA Will Harm Victims of Black Lung." National Health Law Program, February 14, 2017. https://healthlaw.org/wp-content/uploads/2017/02/ACArepealBlackLungSomers-2.15.17.pdf.

Song, Yong, Katherine Southam, B. Basil Beamish, and Graeme R. Zosky. 2022. "Effects of chemical composition on the lung cell response to coal particles: Implications for coal workers' pneumoconiosis." *Respirology* 27, no. 6 (March 20, 2022): 447–54. https://doi.org/10.1111/resp.14246.

Sun, Yingying, Andrew S. Kinsela, and T. David Waite. 2022. "Elucidation of alveolar macrophage cell response to coal dusts: Role of ferroptosis in pathogenesis of coal workers' pneumoconiosis." *Science of The Total Environment* 823 (June 1, 2022): 153727. https://doi.org/10.1016/j.scitotenv.2022.153727.

Toler, Jessica L. 2002. "Dead Canaries: The Struggle of Appalachian Coal Miners to Get Black Lung Benefits Student Note." *Journal of Gender, Race and Justice* 6, no. 1: 163–82. https://heinonline.org/HOL/P?h=hein.journals/jgrj6&i=169.

Torres-Duque, C., D. Maldonado, R. Perez-Padilla, M. Ezzati, G. Viegi, and on behalf of the Forum of International Respiratory Societies (FIRS) Task Force on Health Effects of Biomass Exposure.

2008. "Biomass Fuels and Respiratory Diseases: A Review of the Evidence." *Proceedings of the American Thoracic Society* 5, no. 5 (July 15, 2008): 577–90. https://doi.org/10.1513/pats.200707-100RP.

U.S. Census Bureau. 2023. "American Community Survey." ACS Demographic and Housing Estimates. https://data.census.gov/table?q=DP05:+ACS+DEMOGRAPHIC+AND+HOUSING+ESTIMATES&tid=ACSDP1Y2021.DP05.

U.S. Census Bureau. 2022a. "Growth in U.S. population shows early indication of recovery amid COVID-19 pandemic." US Census Bureau (website). December 22, 2022. https://www.census.gov/newsroom/press-releases/2022/2022-population-estimates.html#:~:text=22%2C%202022%20%E2%80%93%20After%20a%20historically,components%20of%20change%20released%20today.

U.S. Census Bureau. 2022b. "CBP Datasets." County Business Patterns (CBP). https://www.census.gov/programs-surveys/cbp/data/datasets.html.

U.S. Department of Labor. 2023. "Search MSHA Data Sets." Mine Safety and Health Administration (MSHA) Mines Data Set. https://arlweb.msha.gov/OpenGovernmentData/OGIMSHA.asp

U.S. Energy Information Administration. 2023. "Annual Coal Report." https://www.eia.gov/coal/annual/.

West Virginia University, West Virginia & Regional History Center. n.d. "Appalachian Collection." August 24, 2022. https://wvrhc.lib.wvu.edu/collections/appalachian-collection.

Whitney, Eric. 2017. "Native Americans Feel Invisible In U.S. Health Care System." *NPR WAMU 88.5*. December 12, 2017. https://www.npr.org/sections/health-shots/2017/12/12/569910574/native-americans-feel-invisible-in-u-s-health-care-system.

World Health Organization. 2022. "Household Air Pollution." Fact sheets, November 28, 2022. https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health.

Zhang, Junfeng (Jim), and Kirk R Smith. 2003. "Indoor air pollution: a global health concern." *British Medical Bulletin* 68, no. 1 (December 1, 2003): 209–25. https://doi.org/10.1093/bmb/ldg029.

Zhang, Junfeng (Jim), and Kirk R. Smith. 2007. "Household Air Pollution from Coal and Biomass Fuels in China: Measurements, Health Impacts, and Interventions." *Environmental Health Perspectives* 115, no. 6 (June 1, 2007): 848–55. https://doi.org/10.1289/ehp.9479.

Zierold, K. M., A. N. Hagemeyer, and C. G. Sears. 2020. "Health Symptoms among adults living near a coal-burning power plant." *Archives of Environmental & Occupational Health* 75, no. 5 (July 3, 2020): 289–96. https://doi.org/10.1080/19338244.2019.1633992.

Zosky, Graeme R, Ryan F Hoy, Elizabeth J Silverstone, Fraser J Brims, Susan Miles, Anthony R Johnson, Peter G Gibson, and Deborah H Yates. 2016. "Coal workers' pneumoconiosis: an Australian perspective." *Medical Journal of Australia* 204, no. 11: 414–18. https://doi.org/10.5694/mja16.00357.

Zosky, Graeme R., Ellen J. Bennett, Macarena Pavez, and B. Basil Beamish. 2021. "No association between pyrite content and lung cell responses to coal particles." *Scientific Reports* 11, no. 1 (April 14, 2021): 8193. https://doi.org/10.1038/s41598-021-87517-z.

APPENDIX B Quantitative analysis references

B.1: Links to quantitative data sources for coal usage and black lung prevalence statistics

Title	Link
Center for Disease Control and Prevention (CDC) Wide-ranging ONline Data for Epidemiologic Research (WONDER) [1999-2020]	https://wonder.cdc.gov/controller/datarequest/D76
CDC Enhanced Coal Workers' Health Surveillance Program (ECWHSP) [1970-2014]	https://webappa.cdc.gov/ords/cwhsp-database.html
U.S. Census American Community Survey (ACS), 1-year estimates [2014, 2020, 2021]	https://www.census.gov/data/developers/data-sets/acs-1year.html
U.S. Census County Business Patterns (CBP) [1986, 2020]	https://www.census.gov/programs-surveys/cbp/data/datasets.html
U.S. Energy Information Administration (EIA) [1983, 2020]	https://www.eia.gov/coal/annual/
Mine Safety and Health Administration (MSHA) Mines Data Set [1970-2020]	https://arlweb.msha.gov/OpenGovernmentData/OGIMSHA.asp

B.2: Criteria to define Likert scale of coal counties

Scalar	Total Coal Mines (1983)	Total Coal Mines (2020)	Total Coal Production x1000, Short Tons (1983)	Total Coal Production x1000, Short Tons (2020)	% of Residents Employed in Coal Mines (1986)	% of Residents Employed in Coal Mines (2020)	Total Coal Mine Exposure Hours per Week (1970–2022)
No Coal	0	0	0	0	0	0	0
Very low	1	1	1–278.45	1–251.094	0.00001–0.08206	0.00001–0.26296	1–100
Low	2	2	278.46–1,197.262	251.095–1,145.727	0.08207–0.84099	0.26297–0.98682	101–251
Medium	3	3	1,197.263–3,653.943	1,145.728–3,841.668	0.84100–3.55382	0.98683–1.99832	252–518
High	4	4	3,653.944–7,915.196	3,841.669–6,796.696	3.55383–6.74104	1.99833–7.43872	519–1,353
Very high	5+	5+	>7,915.196	>6,796.696	>6.74104	>7.43872	>1,353

Sources: U.S. Energy Information Administration (EIA) 1983 and 2020 collection periods, U.S. Census County Business Patterns (CBP) 1986 and 2020 collection periods, and Mine Safety and Health Administration (MSHA) Mines Data Set 1970-2020 collection period. Note: Similar to the binary definition of coal counties (yes/no), the sliding scale assumes the highest order possible across four criteria (i.e., only the highest metric is used for classification). Distinctions between scalar levels are based on the 25^{th} 50^{th}, 75^{th}, and 90^{th} percentiles in the distribution of data.

B.3: U.S. County coal-producing regions

Region	Count	Percent
Appalachia – North	101	3%
Appalachia – Central	186	6%
Appalachia – South	135	4%
Illinois Basin	17	1%
Interior with Gulf Coast	9	< 1%
Navajo Nation	6	< 1%
Powder River Basin	17	1%
Western	18	1%
Non-Coal	2,647	84%
TOTAL	**3,136**	**100%**

Note: All regions except for Navajo Nation are derived from U.S. Energy Information Administration (EIA)'s coal glossary. Navajo Nation counties were defined according to the U.S. counties that exist either partially or fully inside Navajo Nation borders.

B.3: Correlation matrixes with cumulative black lung cases during the collection period, 1970-2014

Independent Variable	Correlation Coefficient with Black Lung Cases
Total coal mines (1983)	0.6719
Total coal mines (2020)	0.7311
Total coal production (1983)	0.3540
Total coal production (2020)	0.0850
Total coal mining employees (1986)	0.5715
Total coal mining employees (2020)	0.4552
Weekly coal dust exposure hours (1970–2020)	0.7390
Total households with residential coal use (2021)	0.2520

Sources: U.S. Energy Information Administration (EIA) 1983 and 2020 collection periods, U.S. Census County Business Patterns (CBP) 1986 and 2020 collection periods, Mine Safety and Health Administration (MSHA) Mines Data Set 1970-2020 collection period, and U.S. Census Bureau American Community Survey (ACS) 1-year estimates.

B.4: Correlation matrixes with cumulative black lung deaths during the collection period, 1999-2020

Independent Variable	Correlation Coefficient with Black Lung Deaths
Total coal mines (1983)	0.4863
Total coal mines (2020)	0.6289
Total coal production (1983)	0.1776
Total coal production (2020)	0.0380
Total coal mining employees (1986)	0.3639
Total coal mining employees (2020)	0.3123
Weekly coal dust exposure hours (1970–2020)	0.5858
Total households with residential coal use (2021)	0.5339

Sources: U.S. Energy Information Administration (EIA) 1983 and 2020 collection periods, U.S. Census County Business Patterns (CBP) 1986 and 2020 collection periods, Mine Safety and Health Administration (MSHA) Mines Data Set 1970-2020 collection period, and U.S. Census Bureau American Community Survey (ACS) 1-year estimates.

B.5: Correlation matrixes with residential coal use

Independent Variable	Correlation Coefficient with % of county households primarily burning coal for fuel
Total coal mines (1983)	0.2606
Total coal mines (2020)	0.4210
Total coal production (1983)	0.1654
Total coal production (2020)	0.0593
Total coal mining employees (1986)	0.0871
Total coal mining employees (2020)	0.1250
Weekly coal dust exposure hours (1970–2020)	0.3484

APPENDIX C Modeling black lung prevalence

C.1: Analytic Methods

After reviewing descriptive statistics and conducting hypothesis tests, the team wanted to analyze the data a step further by developing a series of statistical models. The purpose of these models to predict the amount of black lung cases or deaths across the United States. Specifically, the team developed a series of least absolute shrinkage and selection operator (LASSO) regressions. LASSO regressions are designed to predict an outcome or dependent variable, such as number of black lung cases in a county, based on a series of independent variables, such as the number of coal mines, share of coal employees, exposure hours, and others. The model assigns weights (coefficients) based on the magnitude of each variable's effect on the outcome (i.e., the more influential, the higher the coefficient in absolute value terms).

The LASSO model is particularly useful in this case because many of the predictor variables are closely related to one another.[29] Tables C.1 and C.2 below show not only the correlations between coal-producing variables, but also the significance of each relationship. The observed correlation between any two of these variables – either in current or former terms— is significant at the 95% confidence level. These matrixes were used to inform the research team's development of the LASSO model.

Table C.1: Covariance matrixes among coal industry indicators, (1983,1986)

	Total mines (1983)	Total Production (1983)	Total employees (1986)
Total mines (1983)	1.0000 -	-	-
Total Production (1983)	0.4471 *0.000*	1.0000 -	- -
Total employees (1986)	0.6990 *0.000*	0.5015 *0.000*	1.0000 -

Sources: U.S. Energy Information Administration (EIA) 1983 collection period and U.S. Census County Business Patterns (CBP) 1986 collection period.

Table C.2: Covariance matrixes among coal industry indicators, (2020)

	Total mines (2020)	Total Production (2020)	Total employees (2020)	Weekly exposure hours (2020)
Total mines (2020)	1.0000 -	-	-	-
Total Production (2020)	0.5391 *0.000*	1.0000 -	- -	-
Total employees (2020)	0.1793 *0.000*	0.7251 *0.000*	1.0000 -	-
Weekly exposure hours (2020)	0.9004 *0.000*	0.7028 *0.000*	0.4126 *0.0000*	1.000 -

Sources: U.S. Energy Information Administration (EIA) 2020 collection period, U.S. Census County Business Patterns (CBP) 2020 collection period, and Mine Safety and Health Administration (MSHA) Mines Data Set 2020 collection period.

[29] In statistics, this phenomenon is commonly referred to as multicollinearity.

The LASSO model penalizes each independent variable such that only the most relevant variables are included in the final estimation process. If two variables are highly correlated, the model will penalize one variable all the way to a coefficient of zero, meaning that it is completely dropped during the estimation process. The LASSO model's final estimation reports all remaining independent variables as well as their respective coefficients. This output results in a formula that estimates how each variable can be used to estimate the number of black lung cases or deaths per county, depending on which outcome variable is used in the model.

During the modeling process, the team tested several potentially influential factors including the geographic region, coal county determinants, comparison of underground to surface mines, and level of residential coal use, among others. Section C.3 discusses the variable selection process and outcomes to the estimations.

C.2: Approach limitations

LASSO models are used to estimate results for a given outcome such as the number of black lung cases or deaths in a county. They do not determine statistical (or causal) significance of independent variables, but rather use the input data to try and predict an outcome. That said, the results of these models are subject to the same data limitations described in the body of this report. Namely, these models are subject to the same constraints of the publicly available data including time frame issues, data suppression, and underreporting. The next section details the findings of these models, keeping in mind these are estimates only and not indicative of perfect information.

C.3: Regression-estimated results

Summit designed four distinct LASSO models where each of the following variables is the dependent variable of interest:

1. Black lung cases;
2. Complicated black lung (PMF) cases;
3. Black lung deaths; and
4. CWP (J60) deaths.

For each model, the team restricted the data to reduce noise and focus on the outcome variables of interest—namely black lung cases and deaths. Each estimate was drawn using counties that report black lung cases or deaths regardless of coal county status. This sub-setting process restricted the input data to roughly 10% of all U.S. counties (293 counties) and estimated results based on observations where black lung disease is most prevalent.[30]

To specify these models, the team used the preceding findings from the summary statistics and hypothesis testing to construct models for black lung cases and deaths. For example, based on the correlation matrixes and principal component analyses, the researchers ensured that the LASSO model retained weekly exposure hours—the variable determined to be one of the most influential on black lung cases and deaths—as part of the final estimation. Other highly correlated variables such as the number of coal workers in 1986 and 2020 may be penalized to zero and dropped by the model. For each

[30] Keeping all 3,136 observations in the dataset would inflate the model's predictive power because the large majority of counties report zero black lung cases or deaths. Restricting the model based on the outcome of interest is designed to reduce noise from these counties.

model, the team used a seed-replicated cross-validation to select the most appropriate lambda with the lowest prediction error and the optimal number of nonzero coefficients.

Additionally, the researchers ensured that county population estimates (from the American Community Survey) also would not be dropped so that the results could be controlled for county size across the country. Given the different data collection periods, the population variable used to control black lung cases or deaths was changed. Since black lung cases were tracked from 1970-2014, models estimating cases incorporate the 2014 population estimates. Conversely, black lung deaths (1999-2020) were estimated using 2020 population data.

Table C.3 below outlines the complete set of independent variables specified for testing in the model.

Table C.3 Black lung modeling: Independent variable specifications

Variable	Use	Required for Estimate?
Population (2014/2020)	Control	Yes
Coal dust exposure per worker (hours)	Predictor	Yes
Number of surface coal mines (1983)	Predictor	No
Number of surface coal mines (2020)	Predictor	No
Number of underground coal mines (1983)	Predictor	No
Number of underground coal mines (2020)	Predictor	No
Number of coal mining employees (1986)	Predictor	No
Number of coal mining employees (2020)	Predictor	No
Surface coal production, x1000 short tons (1983)	Predictor	No
Surface coal production, x1000 short tons (2020)	Predictor	No
Underground coal production, x1000 short tons (1983)	Predictor	No
Underground coal production, x1000 short tons (2020)	Predictor	No
Percent of households with residential coal use	Predictor	No

The following sections highlight the results and key takeaways when modeling black lung cases and deaths in the data.

Modeling black lung cases

When estimating black lung cases, our LASSO model (Table C.4 below) reports an adjusted R^2 value of 0.6833. The R^2 value is a statistic that measures how much variation the model can explain in the data (68.33%). This is considered a high R^2 value based on industry standards and therefore represents a fairly successful model to estimate black lung cases across the United States. Using the same model to estimate severe cases of black lung (PMF) becomes more difficult—this model's adjusted R^2 drops to about 0.5017, which can be found in Table C.5 below. The adjusted R^2 drops when the team estimate PMF cases because the number of PMF cases is less than overall black lung cases (i.e., all severities), making it more difficult to estimate.

Another important statistic for the LASSO model is the mean squared error (MSE). MSE is a measure of how the model's estimates compare to the actual values in the dataset. Simply put, it is a way to see how accurately the model can estimate black lung cases in the data by comparing the estimates to actual results. MSE is the squared amount of the difference between the estimated results and actual data. For this model in Table C.4, the MSE is 471.7127. The square root of 471.7127 is roughly 22, meaning that on average, the LASSO model estimates a result within 22 cases of the actual reported values, whose amounts range from 0 to 876 cases. While researchers always want the lowest MSE possible, this model is considered reasonably powerful in terms of its estimates relative to the actual values reported in the data.[31] Tables C.4 and C.5 below includes the full model output across all coefficients from the model specifications above.

Table C.4: LASSO model estimates: Cumulative black lung cases, 1970–2014

Adjusted (Penalized) R^2: 0.6833 MSE: 471.7127	Observations: 293
Variable	***Coefficient***
Population (2014)	2.56
Coal dust exposure per worker (hours)	67.20
Number of surface coal mines (1983)	0
Number of surface coal mines (2020)	0
Number of underground coal mines (1983)	0
Number of underground coal mines (2020)	4.91
Number of coal mining employees (1986)	0
Number of coal mining employees (2020)	0
Surface coal production, x1000 short tons (1983)	-6.77
Surface coal production, x1000 short tons (2020)	-12.85
Underground coal production, x1000 short tons (1983)	0
Underground coal production, x1000 short tons (2020)	34.07
Percent of households with residential coal use	0

Sources: CDC Enhanced Coal Workers' Health Surveillance Program (ECWHSP) 1970-2014 collection period, U.S. Census Bureau American Community Survey (ACS) 1-year estimates, Mine Safety and Health Administration (MSHA) Mines Data Set 1970-2020 collection period, Energy Information Administration (EIA) 1983 and 2020 collection periods, and U.S. Census County Business Patterns (CBP) 1986 and 2020 collection periods.

The model's coefficient-level output provides information on the level and magnitude of each independent variable's influence on black lung disease. Of the 13 total independent variables subject to inspection, only six were included as part of the final estimation.[32] All variables with coefficients equal to zero are not considered influential enough for our estimation process.

According to the model, exposure to coal dust is the most influential variable for contracting black lung disease along with the amount of coal production in 2020.[33] Interestingly, surface-level coal production is actually estimated have a negative effect on black lung cases in a county. Instead, the team interprets this to mean that surface coal mining has a net-zero impact on the likelihood to contract black lung rather than as a cure to the disease. The key takeaway of this model indicates that coal dust exposure from underground mines is the most powerful predictive characteristic when modeling black lung cases.

[31] LASSO models do not report standard errors or *p* values because they rely on multiple data samples to estimate results.
[32] This includes two variables (population and coal dust exposure) that the research team required the model to include as part of the estimation process.
[33] The number of underground coal mines and coal production in 2020 are highly correlated, suggesting that the number of underground coal mines in 2020 is also an influential factor.

Table C.5: LASSO model estimates: Severe (PMF) black lung cases, 1970–2014

Adjusted (Penalized) R^2: 0. 5026 MSE: 6.1468	Observations: 293
Variable	***Coefficient***
Population	0.18
Coal dust exposure per worker (hours)	6.95
Number of surface coal mines (1983)	0
Number of surface coal mines (2020)	0
Number of underground coal mines (1983)	0
Number of underground coal mines (2020)	0.81
Number of coal mining employees (1986)	0
Number of coal mining employees (2020)	0
Surface coal production, x1000 short tons (1983)	-1.82
Surface coal production, x1000 short tons (2020)	-0.28
Underground coal production, x1000 short tons (1983)	0.40
Underground coal production, x1000 short tons (2020)	0
Percent of households with residential coal use	0

Sources: CDC Enhanced Coal Workers' Health Surveillance Program (ECWHSP) 1970-2014 collection period, U.S. Census Bureau American Community Survey (ACS) 1-year estimates, Mine Safety and Health Administration (MSHA) Mines Data Set 1970-2020 collection period, Energy Information Administration (EIA) 1983 and 2020 collection periods, and U.S. Census County Business Patterns (CBP) 1986 and 2020 collection periods.

Modeling black lung deaths

When modeling black lung deaths, the model has an adjusted R^2 of 0.3749, meaning we can account for about 37% of variability in the data. The MSE is 182.6945, meaning that on average, this model estimates results that fall within about 13 deaths of actual results. For reference, the cumulative death toll in modeled observations range from 0 to 438.

This model is not quite as effective at estimating black lung deaths as cases, but it is still considered an acceptable model to the researchers. Interestingly, if we restrict the model to only estimate CWP (J60) deaths, the adjusted R^2 value slightly improves to 0.3937. This is likely due to the noise from non-CWP deaths (particularly asbestosis) in counties with limited coal influences.

Using the same specifications for this model as the last, the only influential independent variables are those that we specified for the model to use in estimation. Namely, coal dust exposure is the most influential factor in black lung deaths, even controlling for population size. Again, we do not observe residential coal use as a related factor on its own. Tables C.6 and C.7 below present the results of each model estimating black lung and CWP deaths, respectively.

Table C.6 LASSO model estimates: Cumulative black lung deaths, 1999–2020

Adjusted (Penalized) R^2: 0.3749 MSE: 182.6945	Observations: 293
Variable	***Coefficient***
Population (2020)	8.69
Coal dust exposure per worker (hours)	29.85
Number of surface coal mines (1983)	0
Number of surface coal mines (2020)	0
Number of underground coal mines (1983)	0
Number of underground coal mines (2020)	0
Number of coal mining employees (1986)	0
Number of coal mining employees (2020)	0
Surface coal production, x1000 short tons (1983)	0
Surface coal production, x1000 short tons (2020)	0
Underground coal production, x1000 short tons (1983)	0
Underground coal production, x1000 short tons (2020)	0
Percent of households with residential coal use	0

Sources: CDC Wide-ranging ONline Data for Epidemiologic Research (WONDER) 1999-2020 collection period, U.S. Census Bureau American Community Survey (ACS) 1-year estimates, Mine Safety and Health Administration (MSHA) Mines Data Set 1970-2020 collection period, Energy Information Administration (EIA) 1983 and 2020 collection periods, and U.S. Census County Business Patterns (CBP) 1986 and 2020 collection periods.

Table C.7: LASSO model estimates: Black lung CWP (J60) deaths, 1999-2020

Adjusted (Penalized) R^2: 0. 3936 MSE: 101.021	Observations: 293
Variable	***Coefficient***
Population	1.25
Coal dust exposure per worker (hours)	24.28
Number of surface coal mines (1983)	0
Number of surface coal mines (2020)	0
Number of underground coal mines (1983)	0
Number of underground coal mines (2020)	0
Number of coal mining employees (1986)	0
Number of coal mining employees (2020)	0
Surface coal production, x1000 short tons (1983)	0
Surface coal production, x1000 short tons (2020)	0
Underground coal production, x1000 short tons (1983)	0
Underground coal production, x1000 short tons (2020)	0
Percent of households with residential coal use	0

Sources: CDC Wide-ranging ONline Data for Epidemiologic Research (WONDER) 1999-2020 collection period, U.S. Census Bureau American Community Survey (ACS) 1-year estimates, Mine Safety and Health Administration (MSHA) Mines Data Set 1970-2020 collection period, Energy Information Administration (EIA) 1983 and 2020 collection periods, and U.S. Census County Business Patterns (CBP) 1986 and 2020 collection periods.

Estimating underreporting

Between the literature review and discussions with experts, there is reason to believe that some areas, particularly in and around the Navajo Nation, tend to underreport black lung cases and deaths. The reasons for underreporting are discussed more thoroughly in Section 3.3.2.

However, the team can use statistical models to estimate the number of black lung cases and deaths for each county in the Navajo Nation, thereby comparing the reported prevalence of black lung disease to what it may look like using data from the rest of the United States. To test this, Summit designed a series of models using black lung data as well as a series of other independent variables that could be used to estimate black lung prevalence in the Navajo Nation. Among others, these independent variables included the number of coal mines in a county, number of coal employees, exposure to coal dust, and use of coal residentially such as heating or cooking.

Table C.8 below compares the reported number of black lung cases and deaths to Summit's estimated results in the Navajo Nation.

Table C.8: Navajo Nation, reported versus estimated results

County	# Reported Cases	# Estimated Cases	# Reported Deaths	# Estimated Deaths
Apache, AZ	0	1	0	2
Coconino, AZ	0	1	0	4
Navajo, AZ	0	0	0	3
McKinley, NM	0	30	0	20
San Juan, NM	0	22	10	22
San Juan, UT	0	0	0	0

Keeping in mind that counties with 10 or fewer cases are not reported for privacy purposes, the estimated values for these counties reasonably estimate true black lung disease prevalence. While these results are only considered estimates based on a strict set of publicly available data, they may help future researchers and policy makers better understand the disconnect between the prevalence of black lung disease in areas of the Navajo Nation and what is reported in the data.