A bibliometric and visualized analysis of research progress and frontiers on health effects caused by PM$_{2.5}$

Ming Han $^1$ · Fengxia Yang $^1$ · Haifeng Sun $^1$

Received: 30 January 2021 / Accepted: 20 April 2021 / Published online: 28 April 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract
Fine particulate matter (PM$_{2.5}$) is one of the major air pollutants. A large number of epidemiological and experimental studies have shown that PM$_{2.5}$ pollution can cause adverse health consequences, which has attracted more public attention. In order to have a deeper and more structured understanding of the research progress and frontiers on the impact of PM$_{2.5}$ on health, in this study, we used the bibliometrics software CiteSpace to analyze the relevant literature in this field. The results show that since 2000, the relevant literature has increased steadily, especially in the last 5 years, and the number of publications in China has increased rapidly. The United States has the most publications. The Chinese Academy of Sciences and Professor Joel Schwartz are the most published institution and author, respectively, and many articles have been published in the journal of Environmental Health Perspectives. Over time, studies on the health effects of PM$_{2.5}$ have gradually deepened. In addition to a more comprehensive study of its harmful effects, the related molecular mechanisms have also been further explored. We believe that countries and regions should strengthen cooperation and jointly solve the harm caused by PM$_{2.5}$ through the integration of multiple disciplines and fields. In addition, the adverse health consequences and its related mechanisms caused by exposure to ultrafine particle, different chemical components of PM$_{2.5}$, as well as the intervention of the health effects caused by PM$_{2.5}$ need to be further studied.

Keywords  PM$_{2.5}$ · Health effect · Research progress · Web of Science · CiteSpace · Bibliometric

Introduction
With the development of urbanization and industrialization, the problem of air pollution is becoming more and more serious. Although the world has realized this problem and is working hard to improve air quality, air pollution still has a serious adverse impact on public health (Hunter 2020). The World Health Organization (WHO) reported that outdoor and indoor air pollution caused 7 million deaths worldwide in 2016 (Berg and Sturua 2020). PM$_{2.5}$ is one of the main pollutants causing adverse health consequences. According to the Global Burden of Diseases, Injuries, and Risk Factors Study 2019, PM$_{2.5}$ was ranked sixth among level 4 risk factors for deaths, the particulate pollution burden in GBD 2019 was 44.6% higher than in GBD 2017 (Murray et al. 2020). Numerous epidemiological studies have also confirmed that long-term exposure to PM$_{2.5}$ is closely linked to increased mortality from lung and cardiovascular diseases, such as chronic obstructive pulmonary disease (COPD), lung cancer, heart failure, and hypertension (Mannucci et al. 2019; Pope III et al. 2020; Wang et al. 2019; Wang et al. 2018; Zhang and Routledge 2020). With the further in-depth research, it was found that PM$_{2.5}$ exposure can also cause central nervous system diseases, such as Parkinson’s disease (PD), Alzheimer’s disease (AD), as well as adverse birth outcomes (Palacios 2017; Shou et al. 2019; Yuan et al. 2019). Therefore, it has been a hot topic in the relationship between PM$_{2.5}$ exposure and harmful health effects.

Bibliometrics is an effective method that integrates mathematics, statistics, and philology to calculate and study the distribution, structure, quantitative relationship, and laws of document information (Zhang et al. 2020), which has been widely used in the analysis of knowledge structure of many topics (Cheng et al. 2020). Due to the greater public concern
about the adverse health consequences caused by PM$_{2.5}$, there is an increasing number of publications from in vivo, in vitro, and epidemiological studies on the correlation between exposure to PM$_{2.5}$ and the incidence and mortality of various diseases (Turner et al. 2020), as well as the toxic effects and mechanisms. Nevertheless, most of the publications focuses on specific fields are scattered, the interrelationship between them is not yet clear. Although many experts have summarized the results of research in this field (Feng et al. 2016; Kim et al. 2015; Lin et al. 2018; Sharma et al. 2020), there are still few studies on the application of bibliometric methods for analysis. Therefore, it is necessary to explore the development of this field through statistical methods from a macro perspective.

In this study, we searched papers related to the health effects caused by PM$_{2.5}$ from 1990 to 2020 using the Web of Science database and used bibliometric methods to explore, analyze, and summarize the research progress, frontier, and development trend of this field. Hopefully, the results obtained in this study will provide some new ideas for researchers in the field of environmental health science.

### Data and methods

#### Data sources

**Web of Science** We chose the Web of Science Core Collection as the database for obtaining documents, which consists of the Science Citation Index (SCI-E), Social Science Citation Index (SSCI), Arts and Humanities Citation Index (A&HCI), Conference Citation Index (CPCI), and others. It covers a wide range of topics and provides more comprehensive information on literature.

**Retrieval methods** We conducted a comprehensive search in the Web of Science Core Collection database on March 8, 2021, spanning all years, and excluded the literature from 2021 because it was not yet comprehensive enough. The detailed search strategy is shown in Table 1. Considering that PM$_{2.5}$ and health effects are expressed in multiple ways in the papers, we searched and read the titles and abstracts of a large number of related literatures in a rough search before conducting the formal search, and we extracted multiple subject terms that meet the requirements of our subject as thoroughly as possible. In order to ensure the recall rate and the quality of the article, we used topic-specific retrieval. The logical operator “AND” is used to correlate PM$_{2.5}$ and health effects, and the logical operator “OR” is used to correlate different expression of the same concept. Next, we selected the English documents whose document types were articles and reviews and finally got 19,825 documents.

| Table 1 | Retrieval methods and results |
|---------|------------------------------|
| Search Query | Results |
| #1 TS = (PM$_{2.5}$ OR “fine particulate matter” OR “ultrafine particles” OR “nanoparticulate matter” OR “black carbon”) | 58,002 |
| #2 TS = (toxicity OR damage OR injure OR injury OR destroy OR aggravate OR disease OR morbidity OR mortality OR health OR illness OR risk OR “adverse effect” OR “toxicological”) | 19,825 |
| Refined basic document type: (ARTICLE OR REVIEW) AND Time Span = 1985–2020 | |

The use of double quotation marks can help us find phrases more accurately. If you enter a phrase without quotation marks, the search engine will retrieve all records of the entered word.
Bibliometric analysis methods

The main bibliometric methods used in this article are as follows: (1) cooperation network analysis. With the increasing development of science, scientific research cooperation has become an important factor in increasing the results of scientific research, and scientific cooperation is of great significance to information exchange, knowledge dissemination, and resource sharing (Lee and Bozeman 2005). Specifically, if different authors and institutions appear in a paper at the same time, it is considered that there is a cooperative relationship between them. (2) Co-citation analysis. Co-citation analysis in bibliometrics is an analysis method that builds a co-citation network based on the co-occurrence relationship of references, which can be used to show the structure and clustering relationship of citations. If two studies are cited by a later study at the same time, these two studies constitute a co-citation relationship. This concept was coined in 1973 by American intelligence scholar Henry Small (Osareh 1996; Small 1973). The concept of co-citation analysis was first proposed at the references and then expanded to the author and journal, forming the co-citation analysis of authors and journals. (3) Co-occurrence analysis. Co-occurrence of keywords refers to two or more keywords appearing together in a document. Keywords are an important part of the paper. The keywords used in the articles are not only carefully considered by researchers but also intrinsically related to each other. Also, analysis of keywords can indirectly reveal various critical research topics and characteristics of a research field.

The bibliometrics software CiteSpace, developed by Chen (2006), is an application for identifying and displaying new trends and developments in the scientific literature. Its analysis data can come from Web of Science, CNKI, PubMed, and other databases and can visualize the relationships in the literature in the form of scientific knowledge graphs, which not only helps to sort out past research trajectories but also makes predictions on future prospects (Chen et al. 2012). The descriptions of CiteSpace software parameters are shown in Table 2. Our research mainly focuses on the following aspects: (1) analyze the number of publications, that is, use the quantitative visualization report provided by Web of Science to understand the changes in the number of publications each year and the distribution of the number of publications in different countries; (2) understand the important authors, institutions, and journals in the field; (3) dig out the highly influential literature in this field and conducting analysis; (4) understand the research frontiers and hot spots in this field, and (5) explore the future development direction of this field.

In this study, CiteSpace 5.7.R2 software was used to perform cooperation network analysis of authors and institutions, co-citation analysis of journals and references, and co-occurrence analysis of keywords. The document data retrieved from Web of Science were imported into CiteSpace 5.7.R2. In the CiteSpace 5.7.R2 interface, the timespan was set to “1990-2020” (slice length = 1). For the cooperation network analysis, the node type was selected as “author” or “institution,” respectively, and the selection criteria was set to top 30 per slice. For the co-citation analysis, the node type was selected as “cited journal” or “reference,” respectively, and the selection criteria was set to top 15 per slice. For the co-occurrence analysis, the node type was selected as “keyword,” and the selection criteria were set to top 20 per slice (for 1990–2020) and top 50 per slice (for 2016–2020), respectively.

Results and discussion

Publication dynamics

The trend in the number of publications from 1990 to 2020 is showed in Fig. 1. We divided the trend of global total publication into the following three stages: the first stage, from...
1990 to 2000, was characterized by a steady and small number of articles, which did not exceed 10 a year; the second stage, from 2000 to 2010, was shown in a remarkable annual increase in the number of articles published, but the overall growth trend was relatively slow; the last stage was from 2010 to 2020, when the number of articles published was exploded and reached 3389 in 2020. Because the United States, China, and Europe are the three most published...
countries in the world, we analyzed and compared changes in the number of published articles in these three countries. In the first two stages, the trend in the number of published articles in the United States and Europe was roughly the same as in the world, but in the third stage, it grew at a slower rate relative to the world. However, the number of published articles in China was lower until 2010 and began to grow rapidly in 2013. China overtook Europe and the United States in the number of articles published in 2016 and 2017, respectively.

The global distribution of the number of articles published by different countries and their starting dates are shown in Fig. 2. The darker the color of the country on the map, the higher the volume of articles. We found that the United States, China, and many European countries, such as the UK, Germany, Italy, and Spain, have a relatively high numbers of published articles. In addition, India and Japan also have a high volume of publications. In addition, of all countries, the United States was the first country to publish articles (around 1991), while other countries published articles later than the United States, nearly 2000 years (China, the United Kingdom, and Germany all published articles in 1999; Italy and Spain were both in 2000).

We believe that the above results are closely related to the degree of global attention to PM$_{2.5}$. The United States published the article at the earliest time, which is related to its earliest attention to PM$_{2.5}$ pollution. In 1997, the United States first established the Ambient Air Quality Standard (AAQS) for PM$_{2.5}$, which aims to effectively monitor fine particulate matter that has been neglected during industrialization. Some European countries also started the research on PM$_{2.5}$ earlier, mainly on the characteristics, concentration and spatial and temporal distribution of PM$_{2.5}$, and its impact on human health. Later, with the increase of air pollution and the concern of various countries, the number of publications began to increase. By the end of 2010, many countries and organizations in the world had successively issued the PM$_{2.5}$ limit standards (2005 for the WHO; 2008 for the European Union; 2003 for Italy, and 2009 for both Japan and India). As the largest developing country in the world, China’s economy has made prominent achievements in the past decades, which has also brought serious deterioration of air quality. With the increase in the number of haze days, in 2012, China added the PM$_{2.5}$ concentration limit to AAQS. According to the 2013 China Environmental Status Bulletin, China experienced many large-scale haze weathers in 2013, mainly in the central and eastern regions (Wang et al. 2014), when the average number of haze days was the most since 1961, and PM$_{2.5}$ was the primary pollutant.

Analysis of the cooperation network of authors and institutions

The cooperation network of authors and institutions are shown in Figs. 3 and 4, respectively. The size of the nodes in the figure represents the number of articles published by the author or institution. The larger the node or the larger the name font, the more articles the author or institution has published. A link line between authors or institutions represents a cooperation between two authors or institutions, and the thicker the line, the higher the frequency of cooperation. In addition, the
The top 10 authors and institutions by number of published articles are listed in Table 3.

In terms of the author cooperation analysis (Fig. 3 and Table 3), Joel Schwartz (frequency = 209), Petros Koutrakis (frequency = 166), Haidong Kan (frequency = 136), Gerard Hoek (frequency = 134), and Bert Brunekreef (frequency = 126) have larger nodes and name fonts, indicating not only a high number of publications but also strong collaborations with other authors. Professor Joel Schwartz and Petros Koutrakis from Harvard University School of Public Health have published the most articles. Professor Joel Schwartz is interested in epidemiology, and one of his research focuses on acute and chronic effects of air pollutants on health. It has been established that exposure to burning particulates in the air at concentrations well below current standards can have a range of adverse health effects. Professor Joel Schwartz and his team have shown that total mortality, cardiovascular disease, and lung cancer mortality are all positively correlated.

Table 3  Top 10 authors and institutions with the highest frequency in the cooperation network

| Rank | Authors              | Frequency | Rank | Institutions                                      | Frequency |
|------|----------------------|-----------|------|---------------------------------------------------|-----------|
| 1    | Joel Schwartz        | 209       | 1    | Chinese Academy of Sciences                       | 906       |
| 2    | Peter Koutrakis      | 166       | 2    | Harvard University                                | 898       |
| 3    | Haidong Kan          | 136       | 3    | Peking University                                 | 611       |
| 4    | Gerard Hoek          | 134       | 4    | United States Environmental Protection Agency     | 589       |
| 5    | Bert Brunekreef      | 126       | 5    | University of Washington                          | 410       |
| 6    | Brent A Coull        | 120       | 6    | Tsinghua University                               | 377       |
| 7    | Michael Brauer       | 111       | 7    | University of Chinese Academy of Sciences         | 315       |
| 8    | Yang Liu             | 101       | 8    | Fudan University                                  | 303       |
| 9    | Ital Kloog           | 93        | 9    | Utrecht University                                | 283       |
| 10   | Aaron Van Donkelaar  | 88        | 10   | Emory University                                  | 277       |
with ambient PM$_{2.5}$ concentrations (Laden et al. 2006), and exposure to PM$_{2.5}$ can cause cardiac autonomic dysfunction and COPD (Magari et al. 2001; Sunyer et al. 2000). Professor Koutrakis’ research interests focus on human exposure measurement techniques and the sources, transport, and the fate of air pollutants, while other research interests include the assessment of adverse health consequences of the lungs and heart to particulate matter exposure. Professor Koutrakis and his team also used ground-based monitoring data combined with meteorological variables, land-use information, and temporal and spatial lag information to calibrate the chemical transport model (GEOS-Chem) to provide continuous temporal and spatial estimation of air pollutants including PM$_{2.5}$ and its major constituents (Di et al. 2016).

According to the institution cooperation analysis (Fig. 4 and Table 3), Chinese Academy of Sciences has the most publications (frequency = 906), followed by Harvard University (frequency = 898) and Peking University (frequency = 611). The nodes of Harvard University (centrality = 0.26), University of Utrecht (centrality = 0.19), the U.S. Environmental Protection Agency (centrality = 0.18), and University of Washington (centrality = 0.15) have a purple outer ring, which indicates these institutes are more central and play an authoritative role in the cooperative networks. The intensity of the outbursts is marked in red by the inner ring, indicating a high frequency of outbursts (Guo et al. 2020). The nodes of the Sun Yat-sen University (burst = 57.06), University of Southern California (burst = 55.26), and Nanjing University of Information Science & Technology (burst = 44.37) are red, which indicates that these nodes have a sudden increase in the number of publications. In order to understand the specific research directions of the top three institutions with a high volume of publications, we conducted a double search on the basis of Web of Science literature search using each institution as a refined basis and then evaluated their research directions using the Web of Science results analysis tool (Fig. 5). We found that all three institutions have the highest number of publications in the field of environmental sciences. Moreover, Harvard University and Peking University have a high number of publications in the field of public environmental occupational health and toxicology. The Chinese Academy of Sciences and Peking University have a high number of publications in the field of meteorological atmospheric sciences.

**Co-citation analysis of journals and references**

As a quantitative research method in bibliometrics, co-citation analysis of journals can reveal the associations between journals, as well as classify journals. In order to understand the journals that publish the most research papers on the health effects of PM$_{2.5}$ exposure and the collaboration among the
journals, we conducted a journal co-citation analysis. The most frequently cited journals were identified as well. Finally, 70 nodes and 227 connections were obtained, with a density of 0.094. The co-citation network of journals is shown in Fig. 6. The top 10 cited journals, with journal’s citation frequency, centrality strength, burst strength, and IF (2019), are listed in Table 4. The co-citation frequency of each journal is represented by the node size. As shown in Fig. 6, Environmental Health Perspectives had the largest node (frequency = 12,783), followed by Atmospheric Environment (frequency = 11,985) and Environmental Science & Technology (frequency = 10,383). The purple outer ring outside the nodes indicates that these nodes are highly central and play an authoritative role in the field of research. The higher the value, the more important the node. The purple ring of Environmental Health Perspectives (centrality = 0.67), Atmospheric Environment (centrality = 0.39), and Science of The Total Environment (centrality = 0.13) are thicker, indicating that these journals are central and authoritative. We should pay more attention to these journals. Environmental Health Perspectives, the most frequently cited of all journals, is a journal of environmental health research supported by the National Institute of Environmental Health Sciences, which publishes high-quality original research, reviews, and commentaries on experimental toxicology, epidemiology, exposure science, and risk assessment.

According to the co-citation analysis of references, we obtained 205 nodes and 533 links, with a density of 0.0255, which are shown in the co-citation network of references (Fig. 7). The node size indicates the frequency of co-citation. The depth of color represents the year of publication; the cooler the color, the farther away the year, and the warmer the

Table 4 Top 10 cited journals with the highest frequency in the co-citation network of journals

| Rank | Journal                                      | Frequency | Centrality | Burst | IF (2019) |
|------|----------------------------------------------|-----------|------------|-------|-----------|
| 1    | Environmental Health Perspectives            | 12,783    | 0.67       | -     | 8.341     |
| 2    | Atmospheric Environment                      | 11,985    | 0.39       | -     | 4.039     |
| 3    | Environmental Science & Technology          | 10,383    | 0.02       | -     | 7.864     |
| 4    | Science of The Total Environment             | 9564      | 0.13       | -     | 6.551     |
| 5    | Environmental Research                       | 7225      | 0.07       | -     | 5.715     |
| 6    | Journal of The Air& Waste Management        | 6847      | 0.01       | 13.08 | 2.245     |
| 7    | Lancet                                       | 6069      | 0.02       | -     | 60.39     |
| 8    | Environmental Pollution                      | 5565      | 0          | 522.02| 6.793     |
| 9    | Epidemiology                                 | 5554      | 0          | -     | 5.071     |
| 10   | Environment International                    | 5275      | 0          | 423.87| 7.577     |

Fig. 6 The co-citation network of journals
Table 5  Top 10 cited references in the co-citation network of references

| Number of citations | Betweenness centrality | Year of publication | Burst strength | Reference | Title of publication | Cluster |
|---------------------|------------------------|---------------------|---------------|-----------|----------------------|---------|
| 1189                | 0.42                   | 2010                | 333.84        | (Brook et al. 2010) | Particulate matter air pollution and cardiovascular disease | 1       |
| 840                 | 0.25                   | 2012                | 144.04        | (Lim et al. 2012) | A comparative risk assessment of burden of disease and injury attributable to 67 risk factor and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010 | 0       |
| 793                 | 0.11                   | 2017                | 307.7         | (Cohen et al. 2017) | Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015 | 0       |
| 738                 | 0.02                   | 2015                | 197.56        | (Lelieveld et al. 2015) | The contribution of outdoor air pollution sources to premature mortality on a global scale | 0       |
| 590                 | 0.28                   | 2006                | 239.31        | (Pope III and Dockery 2006) | Health effects of fine particulate air pollution: lines that connect | 1       |
| 564                 | 0.12                   | 2014                | 106.96        | (Burnett et al. 2014) | An integrated risk function for estimating the Global Burden of Disease attributable to ambient fine particulate matter exposure | 0       |
| 469                 | 0.08                   | 2013                | 64.81         | (Hoek et al. 2013) | Long-term air pollution exposure and cardio-respiratory mortality: a review | 0       |
| 463                 | 0.12                   | 2002                | 223.99        | (Pope et al. 2002) | Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution | 1       |
| 438                 | 0                      | 2005                | 184.46        | (Oberdorster et al. 2005) | Nanotoxicology: an emerging discipline evolving from studies of ultrafine particles | 0       |
| 407                 | 0.04                   | 2014                | 98.5          | (Huang et al. 2014) | High secondary aerosol contribution to particulate pollution during haze events in China | 0       |

Journal impact factor data comes from the Journal Citation Reports (JCR) in the Web of Science
| Betweenness centrality | Number of citations | Year of publication | Burst strength | Reference | Title of publication                                                                 |
|------------------------|---------------------|---------------------|----------------|-----------|---------------------------------------------------------------------------------------|
| 0.42                   | 1189                | 2010                | 333.84         | (Brook et al. 2010) | Particulate matter air pollution and cardiovascular disease                            |
| 0.34                   | 233                 | 2004                | 106.69         | (Brook et al. 2004) | Air pollution and cardiovascular disease a statement for healthcare professionals from the expert panel on population and prevention science of the American Heart Association |
| 0.34                   | 156                 | 1996                | 93.13          | (Schwartz et al. 1996) | Is daily mortality associated specifically with fine particles?                        |
| 0.34                   | 130                 | 1997                | 74.16          | (Peters et al. 1997) | Respiratory effects are associated with the number of ultrafine particles               |
| 0.31                   | 249                 | 2000                | 133.2          | (Samet et al. 2000) | Fine particulate air pollution and mortality in 20 US Cities, 1987-1994                  |
| 0.28                   | 590                 | 2006                | 239.31         | (Pope III and Dockery 2006) | Health effects of fine particulate air pollution; lines that connect                   |
| 0.28                   | 196                 | 2003                | 93.22          | (Li et al. 2003)   | Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage        |
| 0.26                   | 54                  | 1995                | 32.16          | (Oberdorster et al. 1995) | Association of particulate air-pollution and acute mortality-involvement of ultrafine particles |
| 0.25                   | 840                 | 2012                | 144.04         | (Lim et al. 2012)   | A comparative risk assessment of burden of disease and injury attributable to 67 risk factor and risk factor clusters in 21 regions, 1990-2010; a systematic analysis for the Global Burden of Disease Study 2010 |
| 0.24                   | 98                  | 1994                | 60.15          | (Dockery and Pope 1994) | Acute respiratory effects of particulate air-pollution                                  |
color, the closer the year. Furthermore, the top 10 cited references with high frequency and strong centrality are listed in Table 5 and Table 6, respectively. As shown in Tables 5 and 6, Brook had the most frequency and strong centrality of co-citation (Brook et al. 2010; frequency = 1189; centrality = 0.42).

After careful reading, we found that these literatures focused on different aspects of this research field, and then we briefly divided these research publications into the following categories. The first category, which has the most literature, mainly studied the effects of particulates on human health, with a particular focus on mortality and respiratory and cardiovascular diseases. Particulate matter air pollution and cardiovascular disease (Brook et al. 2010), published by Brook et al. in 2010, is the most frequently cited and highly central literature. This article is an update of the first scientific statement issued by the American Heart Association on “Air pollution and cardiovascular disease” in 2004, which conducted a comprehensive review of the new evidence related to PM exposure and cardiovascular disease and obtained many new conclusions. The main conclusion of this article is that PM$_{2.5}$ exposure, especially long-term exposure, could cause an increase in the mortality of cardiovascular disease and other non-fatal diseases. Pope III and Dockery 2006 investigated lung cancer and cardiovascular mortality under long-term exposure to fine particle pollution and found that long-term exposure to combustion-related fine particle pollution is an important environmental risk factor for cardiovascular and lung cancer mortality (Pope et al. 2002). In addition, Lelieveld et al. (2015), Pope III and Dockery (2006), and Hoek et al. (2013) in Table 5 and Brook et al. (2004), Schwartz et al. (1996), Samet et al. (2000), and Dockery and Pope (1994) in Table 6 published articles are in this category. The second category is the study of the health effects of smaller particulate matter, focusing on ultrafine particles and nanoparticulate matter. Oberdorster et al. (2005) in Table 5 and Peters et al. (1997), Li et al. (2003), and Oberdorster et al. (1995) in Table 6 published articles are in this category. The third category is about the study of the global burden of disease. The literatures published by Lim et al. (2012) and Cohen et al. (2017) in Table 5 belong to this category. The fourth category is about the development of related models and technologies. Burnett et al. (2014) developed an RR model based on the quality of fine particulate matter, which can provide excellent prediction capabilities for the global major death causes of air pollution. The fifth category is about the chemical properties and sources of particulate matter. Huang et al. (2014) investigated the chemical properties and sources of particulate matter.
in some cities in China in 2014, which showed that reducing secondary aerosol emissions is important for controlling China’s PM$_{2.5}$ pollution and reducing the impact of particulate pollution on health, the environment, and the economy.

Based on the above co-citation network analysis of references, we used the keywords as the clustering criteria to perform clustering co-citation with the LLR algorithm. Modularity Q was 0.6784, which is > 0.3, indicating the cluster structure was remarkable. Mean silhouette was 0.6713, which is > 0.5, indicating a convincing clustering. As shown in the references clustering co-citation map (Fig. 8), the research publications were clustered into 7 groups. The size of the node in Fig. 8 indicates the frequency of citation of the reference. Among them, the nodes with the purple outer ring have strong centrality and are of relative importance in the network structure. The thicker the purple outer ring, the stronger the centrality. The digital labels of the clusters are related to the number of documents contained in it. The smaller the number, the greater the number of documents in this category. Publications of #2 sulfates, #3 particles, #4 ambient aerosol, #5 air exchange, and #6 PM$_{2.5}$ mainly focused on the correlation between particulate matter exposure and human health, especially children’s health. These literatures were mostly published in the period 1991–1999, which is the initial stage of this research field. The literatures of #1 heart rate variability were mostly published in 2000–2009, which focused on epidemiological studies of PM$_{2.5}$; #0 nanoparticles has the largest number of publications, most of which have been published since 2010 and have been studied in a wider range of areas. In addition to studying the correlation between nanoparticles and health risks, researchers also focused on the pathogenesis of PM$_{2.5}$ and nanoparticulate matter, the risk of long-term exposure to airborne particulate matter, and the development of functional models of the disease burden caused by particulate exposure.

Co-occurrence analysis of keywords

We used CiteSpace to conduct a co-occurrence analysis of keywords to understand the hot topics and focus about the health effects caused by PM$_{2.5}$. Since space-time elements can better reveal the temporal and spatial evolution process of the research topics, critical for understanding the development of this field. To better reveal the changing trends in research topics over time, we showed time zone distribution of high-frequency keywords in Fig. 9. In 1994, air pollution (frequency = 7051) first entered people’s vision. In the early stage (1993 to 2000), the public paid the highest attention to
death that is the most serious health consequence caused by PM$_{2.5}$, and related keyword was mortality (frequency = 3416, year = 1996). Among all studies related to the diseases caused by PM$_{2.5}$, the correlation between PM$_{2.5}$ and asthma (frequency = 96, year = 1993) and cardiovascular disease (frequency = 299, year = 2000) have the larger number of publications. At this stage, relevant researches also focused on the macro level, mainly explaining the correlation between PM$_{2.5}$ and adverse health effects. With the improvement of the public attention, the research content has become diversified. In the following 10 years, in addition to the deeper research for disease caused by PM$_{2.5}$, the size effects, sources, and mechanisms were also studied. In terms of size effects, studies have shown that the smaller the particle size, the easier it is to adsorb toxic substances (Kwon et al. 2020). According to the prediction of a multi-path particle dose mathematical model developed by the International Radiological Protection Committee, with the decrease of particle size, PM$_{2.5}$ is more likely to be deposited in the lungs and be removed more slowly (Kreyling et al. 2006). With the passage of time, the keywords also changed from PM$_{10}$ (frequency = 734, year = 1999) and PM$_{2.5}$ (frequency = 4718, year = 1999) to nanoparticle (frequency = 298, year = 2007). In terms of sources, different PM$_{2.5}$ components from different sources have different toxic effects, which have attracted much attention from researchers. The relevant keywords are source apportionment (frequency = 1269, year = 2011), emission (frequency = 1354, year = 2002), and polycyclic aromatic hydrocarbon (frequency = 1533, year = 1997).

Regarding the physiological and molecular mechanisms of the health effects of PM$_{2.5}$, oxidative stress (frequency = 1469, year = 1998) and inflammatory injury (frequency = 545, year = 1998) are the two most studied pathogenic mechanisms. In addition, although PM$_{2.5}$ pollution has been improved to some extent with people’s attention, it has not been completely eliminated, and people have been inevitably exposed to PM$_{2.5}$ pollution for a long time. Compared with short-term exposure, long-term exposure (frequency = 835) causes more serious consequence. Since 2009, studies on long-term exposure to PM$_{2.5}$ have been growing. In 2017, the number of publications in China surpassed that in the United States. Since 2017, China has received more and more attention, becoming a keyword in many articles.

In addition, we also conducted co-occurrence analysis of keywords for the last 5 years (2016–2020), which may be more helpful for understanding the hot topics and future directions. Finally, 145 nodes and 297 connections were obtained with a network density of 0.0284. We clustered the keywords using the LLR algorithm, modularity Q was 0.6083, which is > 0.3, indicating that cluster structure was significant. Mean silhouette was 0.8404, which is > 0.5, indicating that the clustering was reasonable. Finally, we obtained eight clusters and the details of each cluster are shown in Table 7. Top 40 keywords with high frequency in keyword co-occurrence analysis (2016–2020) are listed in Table 8.

As shown in Fig. 10, Tables 7 and 8, #0 source apportionment is the first cluster, which indicates researchers’ concern...
about the source and composition of PM$_{2.5}$. In order to effectively reduce PM$_{2.5}$ pollution and its adverse health consequences, it is necessary to have a clear understanding of the source of PM$_{2.5}$. In recent years, more and more researches have focused on the chemical composition characteristics and source analysis of PM$_{2.5}$ in different countries and regions (Ahmad et al. 2021; Gao et al. 2018; Khan et al. 2016; Tan et al. 2017). The term “source apportionment” has also shown high frequency and centrality in the last 5 years (frequency = 1052, centrality = 0.34). Due to the complexity of PM$_{2.5}$

Table 7  Information of keyword clusters (2016–2020)

| Cluster ID | Size | Silhouette Mean (year) | Top terms (log-likelihood ratio, p-level) |
|------------|------|------------------------|----------------------------------------|
| #0         | 26   | 0.925 2016             | source apportionment (119.78, 1.0E-4); PAHs (84.12, 1.0E-4); chemical composition (54.96, 1.0E-4); air pollution (52.34, 1.0E-4); health risk (47.88, 1.0E-4); heavy metals (46.68, 1.0E-4) |
| #1         | 23   | 0.672 2016             | air quality (119.36, 1.0E-4); COVID-19 (80.55, 1.0E-4); ozone (36.44, 1.0E-4); pollution (28.55, 1.0E-4); inflammation (28.12, 1.0E-4); oxidative stress (28.12, 1.0E-4) |
| #2         | 21   | 0.865 2016             | cardiovascular disease (65.59, 1.0E-4); mortality (45.45, 1.0E-4); association (43.69, 1.0E-4); risk (39.85, 1.0E-4); pregnancy (38.94, 1.0E-4); hypertension (36.95, 1.0E-4) |
| #3         | 20   | 0.883 2016             | air pollution (92.93, 1.0E-4); particulate matter (31.97, 1.0E-4); asthma (23.15, 1.0E-4); health (22, 1.0E-4); PM$_{2.5}$ (21.25, 1.0E-4); impact (17.62, 1.0E-4) |
| #4         | 19   | 0.895 2016             | oxidative stress (136.76, 1.0E-4); inflammation (104.57, 1.0E-4); nanoparticles (55.49, 1.0E-4); apoptosis (55.49, 1.0E-4); ultrafine particles (53.73, 1.0E-4); toxicity (33.6, 1.0E-4) |
| #5         | 13   | 0.725 2016             | indoor air quality (92.82, 1.0E-4); personal exposure (32.98, 1.0E-4); air pollution (32.29, 1.0E-4); outdoor (27.56, 1.0E-4); indoor (25.8, 1.0E-4); mortality |
| #6         | 12   | 0.898 2016             | respiratory diseases (23.87, 1.0E-4); short-term exposure (21.02, 1.0E-4); particle size (21.02, 1.0E-4); particulate air pollution (20.47, 1.0E-4); respiratory disease (18.85, 1.0E-4); time-series study |
| #7         | 6    | 0.862 2017             | China (48.23, 1.0E-4); particulate matter (22.63, 1.0E-4); haze (21.81, 1.0E-4); public health (14.99, 0.001); emission inventory (13.04, 0.001); urban sustainability |

The top six terms were shown in each cluster.

Table 8  Top 40 keywords with high frequency in keyword co-occurrence analysis (2016–2020)

| Rank | Keyword                        | Frequency | Centrality | Rank | Keyword                        | Frequency | Centrality |
|------|--------------------------------|-----------|------------|------|--------------------------------|-----------|------------|
| 1    | PM$_{2.5}$                      | 4718      | 0.7        | 21   | PM$_{10}$                     | 734       | 0.05       |
| 2    | air pollution                   | 4325      | 0.53       | 22   | air quality                   | 726       | 0.03       |
| 3    | particulate matter              | 3304      | 0.09       | 23   | inflammation                  | 665       | 0.12       |
| 4    | exposure                        | 2111      | 0.03       | 24   | aerosol                       | 655       | 0.04       |
| 5    | mortality                       | 1866      | 0.15       | 25   | quality                       | 620       | 0.01       |
| 6    | health                          | 1557      | 0.01       | 26   | cardiovascular disease        | 591       | 0.08       |
| 7    | pollution                       | 1262      | 0.12       | 27   | global burden                 | 545       | 0.03       |
| 8    | association                     | 1140      | 0.14       | 28   | pollutant                     | 544       | 0.0         |
| 9    | source apportionment            | 1052      | 0.34       | 29   | model                         | 517       | 0.06       |
| 10   | particle                        | 966       | 0          | 30   | particulate air pollution     | 511       | 0.03       |
| 11   | impact                          | 931       | 0.05       | 31   | ozone                         | 494       | 0.03       |
| 12   | polycyclic aromatic hydrocarbon | 926       | 0.08       | 32   | ambient air pollution         | 491       | 0.08       |
| 13   | emission                        | 913       | 0.07       | 33   | chemical composition          | 470       | 0.01       |
| 14   | ultrafine particle              | 873       | 0.18       | 34   | children                      | 468       | 0.09       |
| 15   | black carbon                    | 870       | 0.03       | 35   | urban                         | 460       | 0.04       |
| 16   | oxidative stress                | 865       | 0.12       | 36   | asthma                        | 442       | 0.02       |
| 17   | disease                         | 840       | 0.06       | 37   | nitrogen dioxide              | 434       | 0.05       |
| 18   | risk                            | 839       | 0.03       | 38   | fine                          | 429       | 0.03       |
| 19   | long term exposure              | 805       | 0.15       | 39   | heavy metal                   | 405       | 0.09       |
| 20   | China                           | 771       | 0.08       | 40   | United States                 | 396       | 0.04       |
sources, its chemical composition (frequency = 470) in the atmosphere is different. Among them, polycyclic aromatic hydrocarbons (PAHs, frequency = 926) have received the most attention, followed by heavy metals (frequency = 405). Studies have shown that the PAHs combined with PM$_{2.5}$ is mainly derived from vehicle emissions, coal combustion, industrial activities, and biomass combustion (Wang et al. 2020). Moreover, PAHs and their derivatives are the most concerned substances among various organic chemical substances related to PM$_{2.5}$, which is also reflected in the health effects (Yan et al. 2019). Keywords under clusters #1 air quality and #3 air pollution include particulate matter and health impact, which are related to our search and selection of data sources on the topic of health impact of PM$_{2.5}$. COVID-19 is a hot topic at present, appearing as a keyword (frequency = 130, year = 2020) in cluster #1. We have studied the relevant literatures and found that whether there is a correlation between air pollution and COVID-19 infection and the role of particulate matter in the mechanism of virus spread has received much attention and discussion (Bontempi 2020; Fattorini and Regoli 2020; Zhu et al. 2020). Cluster #2 cardiovascular disease and #6 respiratory disease show that most current studies on the health effects of PM$_{2.5}$ have focused on cardiovascular diseases (frequency = 591) and respiratory diseases (frequency = 130). Cluster #4 oxidative stress focused on the adverse health consequences caused by particulate matter and its related mechanisms. Ultrafine particles and nanoparticles appeared in this cluster and have a strong correlation, which indicates that the damage caused by particles of smaller size to human body and its mechanism has become a research hotspot; #5 indoor air quality focused on how indoor air pollutants (PM, CO, O$_3$, and VOC) affect human health (Cheek et al. 2021). In the case of individual exposure, researchers pay more attention to the health damage in children caused by indoor air pollution exposure. In previous studies, more attention has been paid to the impact of outdoor air pollution on human health, while the harm of indoor air pollution has been ignored. However, in fact, many people spend more time in indoor environments, such as homes, offices, and schools, coupled with the improvement of people’s quality of life; therefore, the impact of indoor air pollution on human health has gradually attracted attention (Jose et al. 2021); #7 China focuses on the emission inventory of air pollutant in China, urban sustainability, PM$_{2.5}$ concentration and improvement of air quality, as well as the impact of haze and PM$_{2.5}$ pollution on public health.

**Conclusion**

In this current study, based on the Web of Science (WOS) database, visualization software CiteSpace was used to perform a bibliometric analysis on the topic of health effects caused by PM$_{2.5}$. Related publications have been growing steadily since 2000, and publications from China in particular have increased rapidly in the last 5 years. Among all the countries, the United States has the highest number of publications and the earliest publication year. By 2017, the number of published articles from China had surpassed that from the United States. The Chinese Academy of Sciences and Professor Joel Schwartz from the Harvard School of Public Health are the institution and author with the most articles, respectively, and many related articles have been published in the journal of *Environmental Health Perspectives*. Due to differences in the physical properties and chemical composition of PM$_{2.5}$ in various countries and regions, the research results are also inconsistent. Therefore, various countries and regions need to strengthen cooperation to jointly address the adverse effects of PM$_{2.5}$. In addition,
the integration of multiple disciplines and fields can also provide important information for the study of the health effects of PM$_{2.5}$, such as the use of satellite remote sensing and other technologies to estimate the global PM$_{2.5}$ concentration and the development of technologies for the determination and analysis of the chemical composition of PM$_{2.5}$. These studies can provide more accurate PM$_{2.5}$ information and exposure concentration for epidemiological investigations and the experiments in vitro and in vivo, although they are not directly related to our topic. According to the analysis of the changing trends in keywords, we found that the research on PM$_{2.5}$ is gradually deepening, and the harmful effects are also increasing, which involves many molecular mechanisms. In the study of the adverse consequences caused by PM$_{2.5}$ exposure, many experimental techniques and mechanism studies have been relatively mature, especially in the respiratory and cardiovascular system. The analysis of source and chemical composition of PM$_{2.5}$ and the health effects caused by ultrafine particles and nanoparticulate matter have also received attention. However, the adverse health consequences of ultrafine particles and different chemical components in PM$_{2.5}$ and its related toxic mechanisms are still needed to be further explored. At present, PM$_{2.5}$ pollution remains a serious human health hazard. Many countries around the world have taken a series of positive control measures, resulting in a decrease in PM$_{2.5}$ concentration and a significant improvement in air quality. However, PM$_{2.5}$ concentrations in some countries are still above the limits, but even below the limits, its health effects have also changed as its physical and chemical properties have changed. It is related to the characteristics of a country’s economic and social development. Therefore, the research on the health effects of PM$_{2.5}$ is a long-term process. Countries should combine their national conditions and the characteristics of PM$_{2.5}$ pollution to conduct more in-depth studies on the health damage and intervention caused by PM$_{2.5}$.

In this study, scientometric analysis was used. Although this can help us to understand the development process and frontier of the research on the health effects caused by PM$_{2.5}$ in a more structured way, it also has certain limitations. First of all, in this study, we only selected the literature in Web of Science database as the data source, so the selected literature is not comprehensive enough. Secondly, in the process of literature retrieval, in order to ensure the recall rate, we adopt a broader subject retrieval, which inevitably results in the inclusion of some literatures with low relevance. In the subsequent analysis of scientific knowledge maps using CiteSpace, the retrieval data that are not of high relevance will be exposed. Therefore, this study maintains the reliability of the research results to a certain extent, and we suggest that more databases can be combined for a more comprehensive analysis in the subsequent research process.

**Availability of data and materials** All data generated or analyzed during this study are included in this published article.

**Author contribution** Ming Han had the idea for the article and critically revised the work; Fengxia Yang performed the literature search and data analysis and drafted the work, and Haifeng Sun contributed to the language revision of the article. All authors read and approved the final manuscript.

**Funding** This study was supported by the National Natural Science Foundation of China (No. 21707084), Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (201802002), and Natural Science Foundation of Shanxi Province (No. 2019011101).

**Declarations**

**Ethics approval and consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Competing interests** The authors declare no competing interests.

**References**

Ahmad M, Yu Q, Chen J, Cheng S, Qin W, Zhang Y (2021) Chemical characteristics, oxidative potential, and sources of PM$_{2.5}$ in wintertime in Lahore and Peshawar, Pakistan. J Environ Sci 102:148–158. https://doi.org/10.1161/CIR.0b013e3181dbece1

Berg C, Sturua L (2020) The critical need for enhancing the research capacity regarding the intersection of air pollution and non-communicable diseases in Georgia. Georgian Med News 301:178–182

Bontempi E (2020) First data analysis about possible COVID-19 virus airborne diffusion due to air particulate matter (PM): the case of Lombardy (Italy). Environ Res 186:109639. https://doi.org/10.1016/j.envres.2020.109639

Brook RD, Franklin B, Cascio W, Hong Y, Howard G, Lipsett M, Luepker R, Mittleman M, Samet J, Smith SC Jr, Tager I (2004) Air pollution and cardiovascular disease: a statement for healthcare professionals from the expert panel on population and prevention science of the American Heart Association. Circulation 109:2655–2671. https://doi.org/10.1161/01.Cir.0000128587.30041.C8

Brook RD, Rajagopalan S, Pope C et al (2010) Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. Circulation 121:2331–2378. https://doi.org/10.1161/CIRC.0b013e3181dbec1

Burnett RT, Pope C, Arden I et al (2014) An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environ Health Perspect 122:397–403. https://doi.org/10.1289/ehp.1307049

Cheek E, Guercio V, Shrubsole C, Dimitroutlopoulos S (2021) Portable air purification: review of impacts on indoor air quality and health. Sci Total Environ 766:142585–142585. https://doi.org/10.1016/j.scitotenv.2020.142585
Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, Thurston GD (2002) Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. Jama-J Am Med Assoc 287(9):1132–1141. https://doi.org/10.1001/jama.287

Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL (2000) Fine particulate air pollution and mortality in 20 US Cities, 1987-1994. N Engl J Med 343:1742–1749. https://doi.org/10.1056/nejm200012143432401

Schwartz J, Dockery DW, Neas LM (1996) Is daily mortality associated specifically with fine particles. J Air Waste Manage Assoc 46:927–939. https://doi.org/10.1080/10473289.1996.10467528

Sharma S, Chandra M, Kota SH (2020) Health effects associated with PM_{2.5}: a systematic review. Curr Pollut Rep 6:345–367. https://doi.org/10.1007/s40726-020-00155-3

Shou Y, Huang Y, Zhu X, Liu C, Hu Y, Wang H (2019) A review of the possible associations between ambient PM_{2.5} exposures and the development of Alzheimer’s disease. Ecotoxicol Environ Saf 174:344–352. https://doi.org/10.1016/j.ecoenv.2019.02.086

Small H (1973) Co-citation in the scientific literature: a new measure of the relationship between two documents. J Am Soc Inf Sci 24(4):265–269

Sunyer J, Schwartz J, Tobias A, Macfarlane D, Garcia J, Anto JM (2000) Patients with chronic obstructive pulmonary disease are at increased risk of death associated with urban particle air pollution: a case-crossover analysis. Am J Epidemiol 151:50–56. https://doi.org/10.1093/oxfordjournals.aje.a010121

Tan J, Zhang L, Zhou X, Duan J, Li Y, Hu J, He K (2017) Chemical characteristics and source apportionment of PM_{2.5} in Lanzhou, China. Sci Total Environ 601:1743–1752. https://doi.org/10.1016/j.scitotenv.2017.06.050

Turner MC, Andersen ZJ, Baccarelli A, Diver WR, Gapstur SM, Pope CA III, Prada D, Samet J, Thurston G, Cohen A (2020) Outdoor air pollution and cancer: an overview of the current evidence and public health recommendations. CA Cancer J Clin 70:460–479. https://doi.org/10.3322/caac.21632

Wang Y, Yao L, Wang LL, Liu ZR, Ji DS, Tang GQ, Zhang JK, Sun Y, Hu B, Xin JY (2014) Mechanism for the formation of the January 2013 heavy haze pollution episode over central and eastern China. Sci China Earth Sci 57:14–25. https://doi.org/10.1007/s11430-013-4773-4

Wang N, Mengersen K, Kimlin M, Zhou M, Tong S, Fang L, Wang B, Hu W (2018) Lung cancer and particulate pollution: a critical review of spatial and temporal analysis evidence. Environ Res 164:585–596. https://doi.org/10.1016/j.envres.2018.03.034

Wang K, Hao Y, Au W, Christiani DC, Xia Z-L (2019) A systematic review and meta-analysis on short-term particulate matter exposure and chronic obstructive pulmonary disease hospitalizations in China. J Occup Environ Med 61:E112–E124. https://doi.org/10.1097/jjom.0000000000001539

Wang S, Ji Y, Zhao J, Lin Y, Lin Z (2020) Source apportionment and toxicity assessment of PM_{2.5}-bound PAHs in a typical iron-steel industry city in northeast China by PMF-ILCR. Sci Total Environ 713:136428. https://doi.org/10.1016/j.scitotenv.2019.136428

Yan D, Wu S, Zhou S, Tong G, Li F, Wang Y, Li B (2019) Characteristics, sources and health risk assessment of airborne particulate PAHs in Chinese cities: a review. Environ Pollut 248:804–814. https://doi.org/10.1016/j.envpol.2019.02.068

Yuan L, Zhang Y, Gao Y, Tian Y (2019) Maternal fine particulate matter (PM_{2.5}) exposure and adverse birth outcomes: an updated systematic review based on cohort studies. Environ Sci Pollut Res 26:13963–13983. https://doi.org/10.1007/s11356-019-04644-x

Zhang S, Routledge MN (2020) The contribution of PM(2.5) to cardiovascular disease in China. Environ Sci Pollut Res 27:37502–37513. https://doi.org/10.1007/s11356-020-09996-3

Zhang S, Zhao D, Jia W et al (2020) A bibliometric analysis and review of recent researches on TRPM7. Channels (Austin, Tex 14:203–215. https://doi.org/10.1080/19336950.2020.1788355

Zhu Y, Xie J, Huang F, Cao L (2020) Association between short-term exposure to air pollution and COVID-19 infection: evidence from China. Sci Total Environ 727:138704. https://doi.org/10.1016/j.scitotenv.2020.138704

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.