



HAL
open science

Leveraging administrative health databases to address health challenges in farming populations: a scoping review and bibliometric analysis (1975-2024)

Pascal Petit, Nicolas Vuillerme

► To cite this version:

Pascal Petit, Nicolas Vuillerme. Leveraging administrative health databases to address health challenges in farming populations: a scoping review and bibliometric analysis (1975-2024). *JMIR Public Health and Surveillance*, 2025, 11, pp.e62939. 10.2196/62939 . hal-04848292v2

HAL Id: hal-04848292

<https://hal.science/hal-04848292v2>

Submitted on 10 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Original Paper

Leveraging Administrative Health Databases to Address Health Challenges in Farming Populations: Scoping Review and Bibliometric Analysis (1975-2024)

Pascal Petit¹, PhD; Nicolas Vuillerme^{1,2}, PhD, Prof Dr

¹Laboratoire AGEIS, Université Grenoble Alpes, La Tronche Cedex, France

²Institut Universitaire de France, Paris, France

Corresponding Author:

Pascal Petit, PhD

Laboratoire AGEIS

Université Grenoble Alpes

Bâtiment Jean Roget, UFR de Médecine

Domaine de La Merci

La Tronche Cedex, 38706

France

Phone: 33 4 76 63 71 04

Email: pascal.petit@univ-grenoble-alpes.fr

Abstract

Background: Although agricultural health has gained importance, to date, much of the existing research relies on traditional epidemiological approaches that often face limitations related to sample size, geographic scope, temporal coverage, and the range of health events examined. To address these challenges, a complementary approach involves leveraging and reusing data beyond its original purpose. Administrative health databases (AHDs) are increasingly reused in population-based research and digital public health, especially for populations such as farmers, who face distinct environmental risks.

Objective: We aimed to explore the reuse of AHDs in addressing health issues within farming populations by summarizing the current landscape of AHD-based research and identifying key areas of interest, research gaps, and unmet needs.

Methods: We conducted a scoping review and bibliometric analysis using PubMed and Web of Science. Building upon previous reviews of AHD-based public health research, we conducted a comprehensive literature search using 72 terms related to the farming population and AHDs. To identify research hot spots, directions, and gaps, we used keyword frequency, co-occurrence, and thematic mapping. We also explored the bibliometric profile of the farming exposome by mapping keyword co-occurrences between environmental factors and health outcomes.

Results: Between 1975 and April 2024, 296 publications across 118 journals, predominantly from high-income countries, were identified. Nearly one-third of these publications were associated with well-established cohorts, such as Agriculture and Cancer and Agricultural Health Study. The most frequently used AHDs included disease registers (158/296, 53.4%), electronic health records (124/296, 41.9%), insurance claims (106/296, 35.8%), population registers (95/296, 32.1%), and hospital discharge databases (41/296, 13.9%). Fifty (16.9%) of 296 studies involved >1 million participants. Although a broad range of exposure proxies were used, most studies (254/296, 85.8%) relied on broad proxies, which failed to capture the specifics of farming tasks. Research on the farming exposome remains underexplored, with a predominant focus on the specific external exposome, particularly pesticide exposure. A limited range of health events have been examined, primarily cancer, mortality, and injuries.

Conclusions: The increasing use of AHDs holds major potential to advance public health research within farming populations. However, substantial research gaps persist, particularly in low-income regions and among underrepresented farming subgroups, such as women, children, and contingent workers. Emerging issues, including exposure to per- and polyfluoroalkyl substances, biological agents, microbiome, microplastics, and climate change, warrant further research. Major gaps also persist in understanding various health conditions, including cardiovascular, reproductive, ocular, sleep-related, age-related, and autoimmune diseases. Addressing these overlooked areas is essential for comprehending the health risks faced by farming communities and guiding public health policies. Within this context, promoting AHD-based research, in conjunction with other digital data sources (eg, mobile health, social health data, and wearables) and artificial intelligence approaches, represents a promising avenue for future exploration.

KEYWORDS

farming population; digital public health; digital epidemiology; administrative health database; farming exposome; review; bibliometric analysis; data reuse

Introduction

Background

Public health research seeks to identify and understand the factors that influence population health to effectively prevent diseases and promote health and well-being for all [1,2]. A broad range of environmental determinants can impact health across the life span. One of the core areas of public health research, known as the exposome, investigates how cumulative environmental influences contribute to disease etiology and pathogenesis [3-18]. The exposome, which complements genomic research, refers to the comprehensive examination of all environmental exposures experienced throughout an individual's lifetime, including physical, chemical, biological, psychosocial, and behavioral factors, from conception to death [3-18]. The exposome classically includes 3 overlapping domains: the general external exposome (eg, climate and built environment); the specific external exposome (eg, chemical exposure, lifestyle, and occupations); and the internal exposome (eg, aging, oxidative stress, metabolism, and gut microbiome) [8,14,16,17,19,20]. Understanding the exposome is crucial for enabling both population-wide and precision prevention [3,21-23]. However, fully describing the exposome is challenging due to the vast diversity and the temporal and spatial variability of environmental factors [3]. Public health research in this area requires data on both risk factors and adverse health outcomes to progress effectively [3,14,24,25].

The volume of data collected has grown exponentially as the world becomes increasingly reliant on technology and digitization [26,27]. Data are omnipresent in our everyday lives, leading science toward data-driven research [27,28], in particular in the health field. The digital transformation in health care has enabled unprecedented data availability, collection, storage, and analysis capabilities, leading to a paradigm shift in health care systems, with entire care pathways becoming digitized [29,30]. Health-related data now represent approximately 6% of all digital data globally, a figure that continues to rise [31]. This explosion of data has transformed research, providing new opportunities, especially in public health, to enhance disease understanding and evaluate intervention effectiveness [27,28,32-35]. The integration of digital technologies and digital data in public health has led to the emergence of "digital public health," an evolving field focused on using digital data to achieve public health goals [33,36-41]. Public health research is moving from isolated data systems to more integrated, accessible, and reusable data resources [42]. Reusing data allows researchers to explore various health determinants, including environmental, occupational, behavioral, and organizational

factors, fostering a holistic approach to disease prevention and health promotion strategies [14].

Within the digital public health framework, 2 main types of data are being used, namely primary and secondary data. Primary data are tailor-made, designed for a specific purpose, and often used once or repeatedly for the same goal [43-45]. Primary data are the cornerstone of traditional public health policy and decision-making. These data are derived from several types of studies [46-50], in particular observational cohorts (eg, the Framingham Cohort study [51,52]) [46-50,53,54], case-control studies [46-50], cross-sectional surveys (eg, the China Health and Retirement Longitudinal Study [55,56]) [46-50,53], and experimental studies [46-50]. Primary data have many advantages [46-50]. They are rich, of high quality, and are designed to answer specific research questions for public health and epidemiological purposes. Primary data are usually available at the individual level and are derived from studies that control for certain biases. By contrast, they are cumbersome, time-consuming, and costly to set up and maintain [53,54,57]. The representativeness of primary data is also limited in size, geographic scope, and temporal coverage and can erode with time [46-50,53]. Primary data are not free from bias, such as selection, healthy worker, recall, or prevarication biases [53,58].

Unlike traditional public health, digital public health does not rely solely on primary data but takes advantage of the myriad of existing digital data that have not been generated originally for research purposes (ie, secondary data) to overcome some limitations intrinsic to primary data and complement them [28,43,44,53,59-68]. Indeed, some data can have an additional impact when used beyond the context for which they were originally created [68,69]. Secondary data are collected for purposes other than public health or epidemiology and include contextual data (eg, air quality and climate data) [14,24,26,29,70-74], person-generated data (eg, social media, crowdsourcing, and mobile health) [2,24,26,31,43,61,62,73,75-86], synthetic data (eg, digital twin) [87-91], and administrative health databases (AHDs) [26,64,68,81,92-102].

AHD is a broad term encompassing a wide range of routinely collected data on individuals' health and sociodemographic information collected for registration, billing, record keeping, and other administrative purposes [26,64,81,93,95,98,100-102]. For this review, based on previous works [93,95,96,103-105], AHDs included population registers, claims databases, disease registers, electronic health or medical records, and hospital discharge databases that were collected at a local, regional, national, or international level (Table 1) [26,61,62,93,95,96,100,103-110].

Table 1. Definition and characteristics of administrative health databases included in this review.

	Population register	Claims database	Disease register	Electronic health or medical record	Hospital discharge database
Definition	Digital sociodemographic information on the residents of a country	Routinely collected digital information on individual data regarding reimbursement, records of health services, medical procedures, prescriptions, and medical diagnoses	A continuous and exhaustive digital collection of individual data regarding 1 or more health events in a geographically defined population	Systematized digital record of a patient's medical information collected in real time	Digital records of service use with information about patients, their care, and their stay in the hospital
Source	Local or national authorities	Insurance programs or schemes and health care providers	Health care institutions (eg, hospitals)	Hospitals, physicians, health care centers, and institutions	Hospitals
Population	All individuals residing in a country	All individuals covered by an insurance program or scheme	All individuals diagnosed with a specific health event in a population on a geographically defined scale	All patients using the health care system	All patients from a hospital
Purpose or finality	For the administrative purposes of government: to provide reliable information	To store financial and administrative information for medical insurers' and providers' use	For clinical and research purposes: to collect information about people diagnosed with a specific health event	For clinical and billing purposes: to document patients' clinical condition	For billing or accounting purposes
Health event	None	Health events covered by insurance or a health care provider	Specific health events (eg, cancer)	Health events requiring care that are reported in medical records	Health events from hospital admission

AHDs offer many advantages for research. Such data are collected as part of routine administrative processes, reducing additional costs for researchers. Therefore, AHDs offer relatively inexpensive access to a large number of individuals who can be tracked with time for several years, guaranteeing the representativeness of the populations studied [26,54,78,93,95,104,105,111-114]. Data recorded within AHDs are structured, coded in a standardized way, and less affected by participation and recall biases [54,58,95,106,113,115]. AHDs enable the study of rare events and populations underrepresented in studies using only primary data [95,111-113]. AHDs have limitations inherent to their nature, such as the absence of some confounding factors, the limited granularity of certain information, the data complexity, and confidentiality issues [73,78,93,95,115-125].

Rationale

AHDs are increasingly used in population-based health research due to their complementarity with traditional sources of public health and epidemiological data (ie, primary data) [42,59,64,87,93,95,96,126-128]. The reuse of AHDs, referring to their application beyond their original or intended purpose, holds major potential to advance public health and epidemiological research, offering insights that can guide public health decision-making [42,59,64,87,93,95,96,105,126-131]. Although several reviews have previously explored the general use of AHDs in research [42,95,107,129,132-135], others have focused on their application within specific countries [96,136], examined individual AHDs [108,137], or investigated their role in studying specific diseases and adverse health outcomes [44,93,104,138-142]. However, to the best of our knowledge, no study has synthesized how AHDs are reused for

epidemiological and public health research within a specific population group.

To address this gap, we conducted a comprehensive scoping review and bibliometric analysis aimed at identifying how AHDs are used to address health issues in a specific population. We selected farming populations as an illustrative example because they present unique health and disease patterns [143-147]. Globally, approximately 27% of the workforce is engaged in occupational farming, and this group is exposed to numerous risk factors (ie, exposomes), including pesticides, biological agents, and limited access to health care [148]. These exposures put them at heightened risk for a wide range of adverse health outcomes [143,145,147,149]. Although agricultural safety and health have become a major public health issue in recent decades, most research on the health of farming populations has relied on traditional epidemiological and community-based studies, which often face limitations in terms of sample size, geographic scope, temporal coverage, and the range of health events examined [145,150,151].

In this context, AHDs offer valuable opportunities to enhance public health and epidemiological research in farming populations by providing broader insights, identifying at-risk subgroups, and informing health services and policy development [152]. The primary objectives of this scoping review were two-fold: (1) to summarize the current state of AHD-based research in farming populations by examining which types of AHDs are used and why, whether AHDs are integrated with other data sources, which farming populations have been studied, and what exposures and health outcomes have been explored and (2) to identify key areas of interest and potential research gaps and unmet needs in this field.

Methods

Overview

This scoping review was conducted and reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) and evidence maps guidelines (Table S1 in [Multimedia Appendix 1](#)) [153] following a single screening approach. The protocol of this study was not registered. A 7-step procedure was used: research question formulation, identifying relevant publications, title review, abstract review, full-text review, data extraction, and data analysis.

To formulate our research question, we followed the Joanna Briggs Institute guidelines, using the population, concept, and context criteria framework [154]. Our population included all individuals engaged in farming and all individuals exposed to farming-related exposures. The concepts included all possible public health and epidemiological research works that involved the study of a health outcome of interest. The context was the

use, in any setting, of at least one of the AHDs, as defined in [Table 1](#).

Search Strategy and Selection Criteria

To develop and validate the search strategy, previous reviews that examined the reuse of AHDs for population-based research were identified and refined [93,96,103,104]. Our initial search revealed that electronic health records (EHRs) are often interchangeably referred to as electronic medical records (EMRs). A distinction between EHR and EMR is sometimes made, with EMR describing patients' care from only 1 practice (eg, specific encounters in hospitals), which is contrary to EHR [105]. In that case, EMR serves as a data source for EHR. This distinction was not considered in this paper. In addition, to ensure comprehensiveness, the search terms were broadened by searching for their synonyms. For example, search terms such as "electronic health record," "digital health record," "electronic medical record," "EHR," or "EMR" were used as synonyms for electronic health or medical records. A total of 72 terms pertaining to 2 categories (farming and AHDs) were used ([Textbox 1](#)). The search terms were reflective of our research topic and question.

Textbox 1. Search terms.

<p>Farming</p> <ul style="list-style-type: none"> • husbandry* OR agriculture* OR farming OR farm* OR agricultural* OR farmworker* <p>Administrative health databases (combined using AND)</p> <ul style="list-style-type: none"> • "health record" OR "health records" OR "digital record" OR "digital records" OR "health administrative register" OR "health administrative registry" OR "health register" OR "health registry" OR "medical register" OR "medical registry" OR "electronic health record" OR "electronic health records" OR "EHR" OR "EMR" OR "electronic medical record" OR "electronic medical records" OR "digital medical record" OR "digital medical records" OR "digital health record" OR "digital health records" OR "health administrative data" OR "health administrative database" OR "health administrative dataset" OR "health administrative datasets" OR "health administrative databases" OR "administrative health data" OR "administrative health database" OR "administrative health dataset" OR "administrative health datasets" OR "administrative health databases" OR "insurance data" OR "insurance database" OR "insurance databases" OR "insurance dataset" OR "insurance claim" OR "insurance claims" OR "cancer registry" OR "cancer register" OR "health insurance" OR "health surveillance program" OR "health surveillance programs" OR "Mutualite Sociale Agricole" OR "MSA" OR "health insurance system" OR "record-linkage" OR "population register" OR "population registry" OR "insurance scheme" OR "social security scheme" OR "hospital discharge" OR "administrative claim" OR "administrative claims" OR "medical claims" OR "medical claim" OR "electronic claim" OR "electronic claims" OR "mortality register" OR "mortality registry" OR "mortality record" OR "mortality records" OR "disease register" OR "disease registry" OR "illness register" OR "illness registry" OR "disorder register" OR "disorder registry"
--

To develop the eligibility criteria, an initial search of the literature was conducted on PubMed, with a review of the first 100 articles that used AHDs for public health and epidemiological research. In our pilot run, disease and morbidity registers were initially not considered as AHD because they were created for clinical and research purposes [47-50,53,76,155,156]. However, because disease registers contain some information derived from medical records, we

decided to consider them as AHD for this review. The eligibility criteria are presented in [Textbox 2](#). The search was restricted to original peer-reviewed records (all types were included) written in English or French but not constrained by the year of publication [93,106,157]. Publications that examined partly farming populations, with, for instance, studies reporting health risks for various sectors of activity, were included.

Textbox 2. Eligibility criteria for selection of publications.

Inclusion criteria for articles

- Data had to originate at least partly from the administrative health database (AHD)
- The study had to pertain at least partly to the farming population
- The study had to relate to public health or epidemiological research
- Original peer-reviewed publications
- Publications in English or French

Exclusion criteria for articles

- Publications not describing the use of an AHD
- Animal or in vitro studies
- Publications not in English or French

The final literature search was conducted on both PubMed and Web of Science Core Collection databases. Regarding the Web of Science Core Collection, a topic search was performed. To reduce the bias induced by daily database changes, all data collection (literature retrieval and data download) was conducted and completed on the same day, that is, April 15, 2024. Titles, abstracts, and full-text publications were screened based on pre-established inclusion and exclusion criteria. The inclusion criteria for each phase of the literature search are provided in Table S1 in [Multimedia Appendix 1](#). When abstracts did not contain enough information about correspondence to inclusion or exclusion criteria, the article was considered for full-text review. Reference lists of included publications were not searched, although they might have also yielded new relevant studies.

Data Collection and Processing

A total of 29 metadata were extracted from each publication included in the scoping review ([Table 2](#)).

The data underwent rigorous manual validation, cleaning, and harmonization following a structured 5-step process. First, duplicate items (eg, keywords and institutions) were removed. Second, leading and trailing white spaces were eliminated. Third, items were standardized by converting text to lower case, with only the first letter capitalized. In the fourth step, items were harmonized to either singular or plural forms consistently. Finally, synonyms or terms with similar meanings (eg, “illness” and “disease”) were unified under a single term. For instance, “Pesticide,” “Pesticide exposure,” and “Pesticide use” were standardized to “Pesticide,” while “Pulmonary disease copd,” “Copd,” and “Chronic obstructive pulmonary disease” were unified as “COPD.” For cancer-related keywords, the International Classification of Diseases, eleventh revision, was used to consolidate varied terms (eg, “lung cancer,” “lung cancer risk,” “lung and bronchus cancer,” “lung tumor,” “lung tumour,” “lung neoplasm,” and “basal cell carcinoma of the lung”) into standard categories (eg, lung cancer). Quality appraisals were not performed because they were beyond the aim of this review [106,157].

Table 2. List of metadata of interest to collect from the literature search.

Metadata	Fictional example
Publication year	2024
Publication type	Article
Study name	Project X
Goal of the study	To study the association between farming and health outcome
Study type	Ecological study
Is the study nationwide?	Yes
Digital data used	Insurance claims
Goal of the digital data used	To identify farmers
Is active data used?	Yes
Active data used	Clinical examination
Farming exposure considered	Farming activity and pesticide compounds
Farming activities studied, n	10
Pesticide compounds studied, n	29
Population	Adults
Sex	Female
Participants included, n	100 to 1000
Country	France
Oldest data used (year)	1991
Most recent data used (year)	2020
Data follow-up period (years)	4
Years between the most recent data used and publication year, n	7
Disease or health events	Parkinson disease
Authors' names	Gauthier J
Authors' keywords	Pesticide
Authors' country	France
Authors' institution	Université Grenoble Alpes
Journal	Environmental Health Perspectives
Funding body	MIAI@Grenoble Alpes ^a
Citations, n	14

^aMIAI@Grenoble Alpes: Multidisciplinary Institute in Artificial Intelligence at the Université Grenoble Alpes.

Data Analysis

To analyze the research directions (ie, hot spots and gaps) on the use of AHDs for public health and epidemiological research in farming populations, a bibliometric approach was conducted [158-160]. This analysis examined the number of publications, countries of publications, most active journals, institutions, authors, funding bodies, subject areas, citations of publications, and keywords of publications. Seven bibliometric metrics were computed, including the h-, g-, m-, and Y-indices; dominance factor; annual growth rate (AGR); and fractionalized frequency (Table S3 in [Multimedia Appendix 1](#)). The h-index attempts to measure both the productivity and citation impact of the published body of work of an entity (eg, author, institution, and

journal) [161,162]. It refers to the total number of publications by a particular entity with at least the same number of citations. The m-index is calculated by dividing the h-index by the number of years of an entity's productive life (eg, researcher) [161]. The g-index of an entity corresponds to the largest number g such that the top g publications have at least $\geq g^2$ citations together [162]. The Y-index refers to the sum of both the total number of first-authored publications and the total number of corresponding author publications [163]. The dominance factor refers, for a particular researcher, to the proportion of multiauthored publications as a specific author's rank to the total number of multiauthored publications [164]. The fractionalized frequency intends to reflect an author's contribution. The AGR refers to the variable's change in

percentage as a year-over-year statistic [165]. The most up-to-date journals' impact factors and ranks were retrieved manually using the Journal Citation Report in April 2024.

Spearman correlations were calculated to examine the association between the number of publications and gross domestic product (GDP); population size [166]; and the total labor force, the number of researchers in research and development (per million people), fertilizer consumption (in both % of fertilizer production and kilograms per hectare of arable land), agricultural land (km²), agricultural land (% of land area), land under cereal production (hectares), permanent cropland (% of land area), cereal production (metric tons), crop production index, food production index, livestock production index, cereal yield (kilogram per hectare), female individual employment in agriculture (% of female employment), male individual employment in agriculture (% of male employment), employment in agriculture (% of total employment); and agriculture, forestry, and fishing, value added (% of GDP). These country characteristics were obtained from the World Bank. The most recent country characteristic (eg, GDP) was used when available.

Research directions, including hot spots and gaps, were investigated with keyword frequency, co-occurrence (counting of paired keywords), and thematic mapping analyses. Thematic mapping and keyword co-occurrence network are 2 complementary but distinct approaches that serve different purposes and offer different insights. In summary, thematic mapping focuses more on the strategic positioning of research themes within a field, while keyword co-occurrence networks emphasize the relationships and connections between specific keywords in the literature [158,167]. Both methods complement each other and are usually used to provide a more comprehensive understanding of research landscapes. The co-occurrence of 2 keywords was defined by the frequency with which they appear together in publications and was quantified using association strength (AS) or equivalent index, calculated as $AS_{ij} = \frac{c_{ij}}{c_i c_j}$, where c_{ij} is the number of publications in which keywords i and j co-occur, while c_i and c_j are the number of publications in which each keyword appears, respectively [158,167]. AS measures how close 2 keywords are to each other. An AS value of 1 indicates keywords always appear together, while 0 indicates they never co-occur. These keyword co-occurrences can be visualized using a co-occurrence network graph, where a vertex or node represents a keyword, the size of the node represents the keyword frequency, and the edge represents the association between 2 keywords [158,167]. On the basis of the keyword co-occurrence network graph, a community detection procedure can be used to identify groups of words highly associated with each other [158,167]. In other

words, equivalent keywords based on AS can be grouped together to identify research themes [158,167]. A strategic diagram or thematic map is based on Callon centrality (x-axis) and Callon density (y-axis) [158,167]. Callon centrality measures the degree of interaction of a theme with other themes.

It is defined as $Callon\ centrality = 10 \times \sum AS_{kh}$, where k is a keyword belonging to a theme and h is a keyword belonging to another theme [158,167]. Callon centrality can be interpreted as an indicator of the importance of a particular topic within the broader research landscape. Callon density measures the internal strength of a theme. It is defined as $Callon\ density = 100 \times \sum \frac{AS_{ij}}{w}$, where i and j are keywords belonging to the same theme and w is the total number of keywords in a theme [158,167]. Callon density serves as a metric for assessing the progression and maturation of that topic [158,167]. A strategic diagram is divided into 4 quadrants according to Callon centrality and density values, which correspond to 4 types of topics. Hot spots or hot topics are defined by both high density and high centrality values (upper-right quadrant), while basic topics are defined by high centrality but low density values (lower-right quadrant). Peripheral topics are defined by both low centrality and low density values (lower-left quadrant), while niche topics are defined by low centrality and high density values (upper-left quadrant) [158,167].

To focus on agricultural or farming exposome research, a bibliometric profile of the “farming exposome” was constructed, which restricts the exposome concept to environmental exposures specific to farming populations [152,168]. This bibliometric farming exposome picture examined co-occurrences between keywords related to potential risk factors and specific health events (eg, cancers and reproductive disorders).

The bibliometric analysis was conducted and reported according to the preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO; Table S4 in [Multimedia Appendix 1](#)) [169]. All analyses were performed using R software (version 4.3.2; R Foundation for Statistical Computing) for Windows 10 (Microsoft Corporation). The bibliometric analysis was performed using the *bibliometrix* R package (version 4.1.4) [170].

Results

Overview

After excluding 4485 irrelevant records, 296 publications were analyzed ([Figure 1](#)). The majority were articles (293/296, 98.9%), with a small number of reviews (2/296, 0.7%) and editorial materials (1/296, 0.3%; [Table 3](#)). Only one-third of the publications (107/296, 36.1%) were open access ([Table 3](#) and [Figure S1](#) in [Multimedia Appendix 1](#)).

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flowchart depicting the literature search and the evaluation process for finding relevant records. The search, conducted on April 15, 2024, in PubMed and Web of Science, had no date restrictions. AHD: administrative health database.

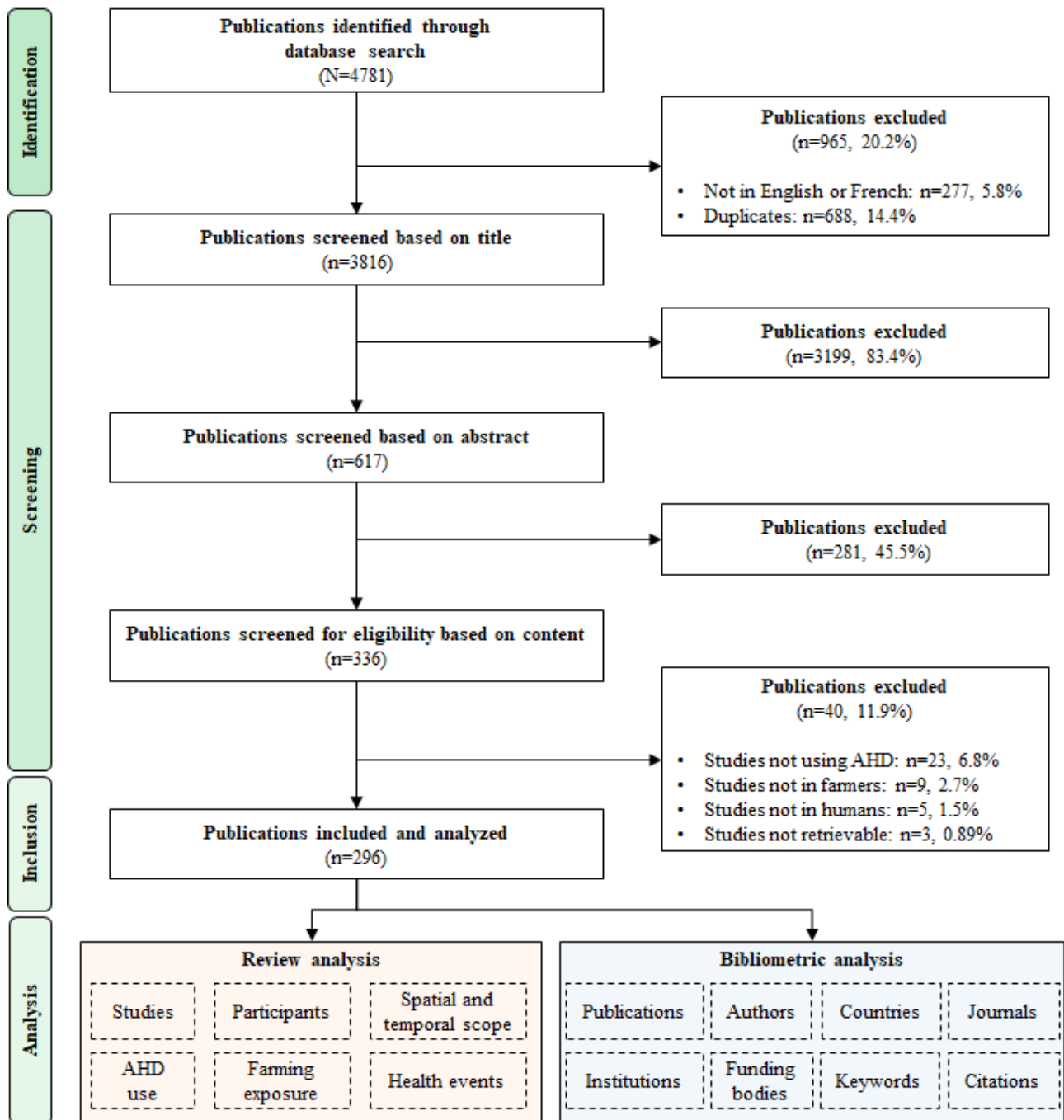


Table 3. Main characteristics of the included publications (N=296).

Description	Results
Timespan	1975 to 2024
Publication type, n (%)	
Article	293 (99)
Review	2 (0.7)
Editorial material	1 (0.3)
Open-access publications, n (%)	107 (36.1)
Document age (y), mean (SD)	14.2 (11.8)
Annual growth rate (%)	5.2
Publication citations	
Total, n ^a	9379
Average citations per publication	31.7
Average citations per year per publications	2.02
References, n	8814
Journals	
Total, n	118
Average number of publications per journal	1.86
Average number of citations per journal	79.5
Authors	
Total, n	1225
Single-author publications, n	4
Author appearances, n	1882
Average number of coauthors per publication	6.36
Average number of publications per author	0.24
International coauthorships (%)	24.3
Author's keywords, n	576
Author's country	
Total, n	34
Average number of publications per country	2.86
Average number of citations per country	436.0
Author's institution	
Total, n	338
Average number of publications per institution	3.11
Average number of citations per institution	101.3
Author's funding body	
Total, n	181
Average number of publications per funding body	2.48
Average number of citations per funding body	77.7

^aTotal, n indicates that the respective parameter has been cited n number of times, as in 296 publications have been cited 9379 times.

The average publication age was 14.2 (SD 11.8) years, ranging from the oldest in 1975 [171] to the most recent in April 2024 [152]. From 1975 onward, there has been a steady increase in publications using AHDs to address health issues in farming

populations, with an AGR of 5.2%. Notably, almost one-third of these articles (91/296, 30.7%) were published in the last 5 years, highlighting the rising interest in AHD-based public health research in this population (Figure S2 in [Multimedia](#)

[Appendix 1](#)). Collectively, the publications received 9379 citations, averaging 31.7 citations per publication. Figure S3 in [Multimedia Appendix 1](#) presents the historical direct citation network. The body of work involved 1225 authors from 338 institutions, with 1882 author appearances and an average of 6 authors per paper ([Table 3](#)). Four (1.4%) out of the 296 publications were single-author publications. On average, each paper cited 30 references.

Studies were led by authors from 34 countries, predominantly high-income nations, with 24.3% (72/296) of studies involving multicountry collaborations (Figure S4 in [Multimedia Appendix 1](#)). US-based authors contributed the most publications (91/296, 30.7%), followed by authors based in France (71/296, 24%) and Finland (35/296, 11.8%). US authors also had the most citations (3495/9379, 37.2%), with France and Finland ranking second and third, respectively.

Of 296 publications, the 25 (8.4%) most cited ones, appearing in 17 different journals, received between 83 (83/9379, 0.9%) and 485 (485/9379, 5.2%) citations ([Table S5 in Multimedia Appendix 1](#)) [[150,172-196](#)]. Of these 25 publications, 10 (40%) were published before 2000, another 10 (40%) between 2000 and 2010, and 5 (20%) after 2010. Most of these studies focused on cancer risk (16/25, 64%), while others investigated neurodegenerative disorders (5/25, 20%); respiratory conditions (2/25, 8%); and multiple health outcomes, such as sleep disorders, mental health disorders, and musculoskeletal disorders (2/25, 8%).

[Table S6 in Multimedia Appendix 1](#) provides details on the most productive countries, prolific authors, active journals, institutions, and funding bodies.

Study Characteristics

[Table 4](#) provides an overview of the included publications. Longitudinal study designs were the most common, including

retrospective cohorts (129/296, 43.6%) and prospective cohorts (56/296, 18.9%). Case-control studies (62/296, 20.9%), cross-sectional studies (39/296, 13.2%), and ecological studies (17/296, 5.7%) were less common ([Multimedia Appendix 2](#)). A few studies (10/296, 3.4%) used multiple study designs [[188,194,197-204](#)].

The median follow-up period was 9.5 (IQR 5-17) years. On average, there was a 7-year gap (90% CI 3-14) between the most recent data used and the year of publication, with considerable variation depending on the publication year ([Figure 2](#)). The oldest data were from 1801 [[205](#)], and the most recent data were from 2022 [[206](#)]. Notably, one-third of the data used (98/296, 33.1%) were from before 2000, while nearly three-quarters (214/296, 72.3%) were from before 2015 (Figure S2 in [Multimedia Appendix 1](#)). Of 296 studies, only 10 (3.4%) used data from the last 5 years (from 2020), while 80 (27%) used data from the last 10 years (from 2015).

Studies were conducted in all continents, but most participants were from Europe (249/296, 84.1%), followed by North America (85/296, 28.7%), Asia (24/296, 8.1%), Oceania (17/296, 5.7%), Africa (4/296, 1.4%), and Central and South America (4/296, 1.4%). France (70/296, 23.6%) and the United States (67/296, 22.6%) were the most represented countries, followed by Finland (36/296, 12.2%), Sweden (32/296, 10.8%), Denmark (28/296, 9.5%), and Norway (26/296, 8.8%; [Figure 3](#) and [Figure S5 in Multimedia Appendix 1](#)). Most studies had a regional or local scope (177/296, 59.8%), in particular, traditional epidemiological studies, such as Agriculture and Cancer (AGRICAN) [[207](#)] and Agricultural Health Study (AHS) [[172](#)], which used AHDs to either identify potential individuals for inclusion or enrich their cohorts.

Table 4. Characteristics of the included studies (1975 to 2024; N=296).

Characteristic	Values
Research goal, n (%)	
Study the association between farming and a health event	156 (52.7)
Study the association between individual characteristics and a health event	131 (44.3)
Other research goals	9 (3)
Study design, n (%)	
Retrospective cohort	129 (43.6)
Case-control study	62 (20.9)
Prospective cohort	56 (18.9)
Cross-sectional study	39 (13.2)
Ecological study	17 (5.7)
Multiple designs	10 (3.4)
Review	2 (0.7)
Perspective	1 (0.3)
Geographic scope, n (%)	
Nationwide	117 (39.5)
Regional or local	176 (59.5)
Temporal scope (y)	
Follow-up period, median (IQR)	9.50 (5-17)
Follow-up period, mean (SD)	12.8 (14.0)
Gap between the latest data used and publication year, median (IQR)	7.21 (5-9)
Gap between the latest data used and publication year, mean (SD)	7.21 (4.67)
Population, n (%)	
Adult	265 (89.5)
Adult and child	19 (6.4)
Child	8 (2.7)
Not reported	1 (0.3)
Sex, n (%)	
Female	130 (43.9)
Male	169 (57.1)
Female and male	188 (63.5)
Not specified	108 (36.5)
Participants, n (%)	
>1,000,000	50 (16.9)
100,001 to 1,000,000	53 (17.9)
10,001 to 100,000	65 (22)
1001 to 10,000	67 (22.6)
101 to 1000	47 (15.9)
10 to 100	8 (2.7)
Not reported	3 (1)
AHD^a type, n (%)	
Disease register	158 (53.4)
Electronic health or medical record	124 (41.9)

Characteristic	Values
Insurance claim	106 (35.8)
Population register	95 (32.1)
Hospital discharge databases	41 (13.9)
AHD use, n (%)	
Obtain information on sociodemographics	272 (91.9)
Obtain information on a health event	269 (90.9)
Identify a farmer	147 (49.7)
Identify an individual	140 (47.3)
Obtain information on occupations	117 (39.5)
Exposure assessment	57 (19.3)
Obtain information on a farming activity	43 (14.5)
Other uses	14 (4.7)

^aAHD: administrative health database.

Figure 2. Number of years between the most recent data used and publication for all included articles (1975-2024). Points refer to the average number of years or gap between the most recent data used and publication (x-axis) for each publication year (y-axis). Error bars refer to the 90% CI of the number of years between the most recent data used and publication.

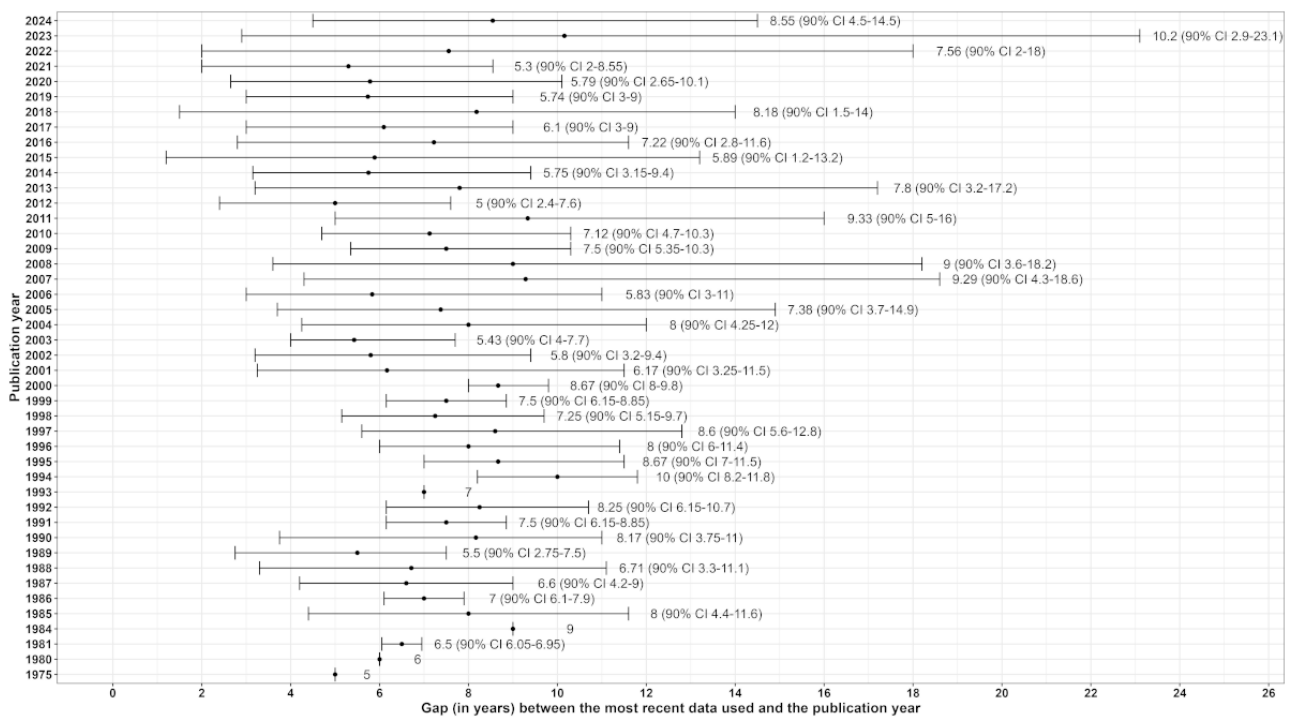
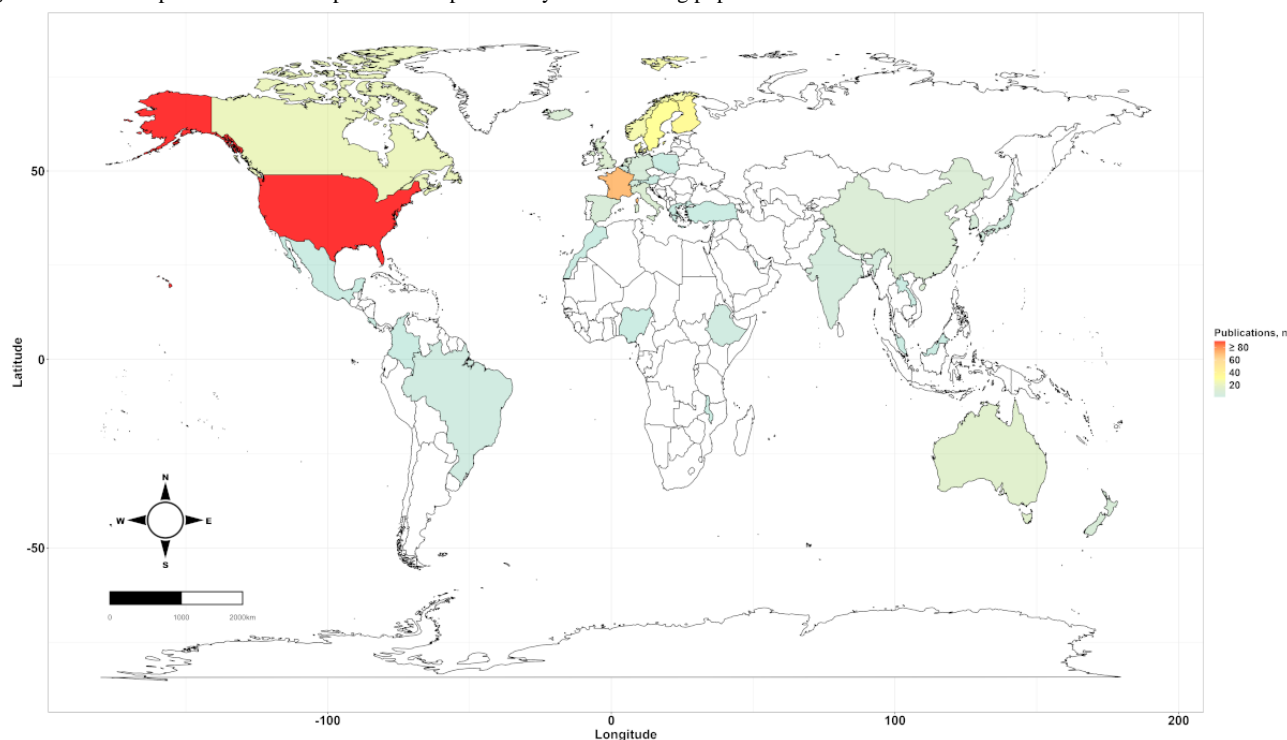


Figure 3. World map of the number of publications per country of the farming population studied between 1975 and 2024.

Most studies included 1001 to 10,000 participants (67/296, 22.6%), followed by studies with 10,001 to 100,000 participants (65/296, 22%) and 100,001 to 1 million participants (53/296, 17.9%; [Table 4](#)). Larger studies (>1 million participants) accounted for 16.9% (50/296) of the included publications. Smaller studies, with 100 to 1000 participants, were less common (47/296, 15.9%), and very few (8/296, 2.7%) had <100 participants. Most studies included adult participants (284/296, 95.9%). Of 296 studies, 169 (57.1%) examined male individuals, 130 (43.9%) examined female individuals, and 188 (63.5%) examined both sexes, but 108 (36.5%) did not specify the participants' sex.

More than half of the studies (156/296, 52.7%) aimed to explore the relationship between farming activities (eg, dairy farming) and health events, while 131 studies (44.3%) focused on individual characteristics, such as occupation, age, sex, and socioeconomic status (Figure S6 in [Multimedia Appendix 1](#)). Among those studies examining individual characteristics, farming was often considered broadly and compared to other occupations (95/131, 72.5%). Conversely, in studies investigating health outcomes specifically related to farming activities, agriculture was treated as a broad category in only 27.6% (43/156) of the cases. Most studies (277/296, 93.6%) used the general population or other nonfarming groups as the reference category without differentiating farmers by job role (eg, farm managers vs farm workers). Descriptive statistics and multivariable regression were the most commonly used methods. Notably, only 2 studies (2/296, 0.7%) incorporated artificial intelligence (AI) in their analysis [[208,209](#)].

Few studies investigated health outcomes in farmers' family members or nonfarmers exposed to farming. Of 296 studies, only 3 (1%) focused on health events in farmers' partners [[177,210,211](#)], 5 (1.7%) on farmers' children [[179,212-215](#),

and 6 (2%) on nonfarmers exposed to farming-related risks [[209,216-221](#)]. There were 11 (3.7%) studies that explored health risks in migrant workers.

Some publications reported findings from the same cohorts (Figure S7 in [Multimedia Appendix 1](#)). The 10 most prolific cohorts included France-based AGRICAN (18/296, 6.1%) [[207](#)], the US-based AHS (17/296, 5.7%) [[172](#)], Nordic Occupational Cancer Study (NOCCA; 12/296, 4.1%) from Nordic countries (Finland, Denmark, Norway, Sweden, and Iceland) [[189](#)], France-based Tracking and Monitoring Occupational Risks in Agriculture (TRACTOR; 7/296, 2.4%) [[222](#)], and Cancer in the Norwegian Agricultural Population (7/296, 2.4%) [[182](#)] cohorts. Other notable cohorts included the US-based United Farm Workers of America (6/296, 2.0%) [[223](#)], France-based BALISTIC (5/296, 1.7%) [[224](#)], the international (29 countries) consortium agricultural cohort (AGRICOH; 4/296, 1.4%) [[150,225](#)], AIRBAg (4/296, 1.4%) from France [[226](#)], and the US-based National Agricultural Workers Survey (3/296, 1%) [[227](#)]. Among these top 10 cohorts, only NOCCA, United Farm Workers of America, and TRACTOR exclusively used AHDs.

AHD Use

There was high heterogeneity in the coding systems used and the granularity of the information available regarding health events (outcomes), population, and exposure determinants, depending on the AHD and study considered. Regardless of the publication reviewed, AHDs and other datasets were never reported as adhering to the findable, accessible, interoperable, and reusable (FAIR) data principles [[228-230](#)]. In addition, none of them could be considered as FAIR data because, with a few exceptions [[222](#)], most AHDs were not precisely described, and data availability statements were rare. Furthermore, mainly due to privacy concerns, AHDs were not available for open and free access.

The most commonly used AHDs were disease registers, used in more than half of the studies (158/296, 53.4%), followed by electronic health or medical records (124/296, 41.9%), insurance claims (106/296, 35.8%), population registers (95/296, 32.1%), and hospital discharge databases (41/296, 13.9%; [Table 4](#)). Among disease registers, cancer (120/158, 75.9) and mortality registers (75/158, 47.5%) were the most frequently used (Figure S8 in [Multimedia Appendix 1](#) and [Multimedia Appendix 2](#)). Nearly one-third of the studies (91/296, 30.7%) relied on a single AHD, with disease registers being the most common (38/91, 42%), followed by insurance claims (29/91, 32%), electronic health or medical records (18/91, 20%), population registers (5/91, 5%), and hospital discharge databases (1/91, 1%). Other types of digital data were used less frequently, including pesticide registration records (13/296, 4.4%), job-exposure matrices (JEMs; 12/296, 4.1%), crop-exposure matrices (11/296, 3.7%), pesticide use records (8/296, 2.7%), climate data (7/296, 2.4%), and air quality data (2/296, 0.7%). While contextual data were sometimes used (9/296, 3.0%), person-generated data, smart agriculture data, and omics were never used.

The AHDs and other digital data were primarily used to obtain sociodemographic information (272/296, 91.9%) and health event data (269/296, 90.9%). They were also used to identify farmers (147/296, 49.7%) or individuals (140/296, 47.3%), gather occupational information (117/296, 39.5%), assess exposure (57/296, 19.3%), obtain data on farming activities (43/296, 14.5%), or track climate conditions (7/296, 2.4%).

Nearly two-thirds of the studies (181/296, 61.1%) relied exclusively on digital data (AHDs or other), while more than one-third (112/296, 37.8%) incorporated self-reported information/active data (requiring active participant involvement) as part of epidemiological cohorts. A total of 111 (37.5%) out of 296 studies used participant-completed questionnaires (paper or electronic) to gather sociodemographic data and confounding factors (98/296, 33.1%), assess exposure (96/296, 32.4%), or collect health information (83/296, 28%). Some information was obtained through interviews (44/296, 14.9%) or clinical examinations (32/296, 10.8%). Biological monitoring (24/296, 8.1%) and airborne monitoring (2/296, 0.7%) were sometimes used, whereas no study reported dermal monitoring (Figure S9 in [Multimedia Appendix 1](#)).

Among all the AHDs used, the Mutualité Sociale Agricole (MSA) is a singularity. To the best of our knowledge, it is the only AHD specifically dedicated to the entire farming population of a country. Indeed, MSA is the French national insurance scheme that covers the entire farming workforce (5% of the overall French population) [[115,128](#)]. MSA was used in 60 studies (60/296, 20.3%). These studies were often part of cohorts with multiple publications, such as AGRICAN (18/60, 30%), TRACTOR (7/60, 12%), BALISTIC (5/60, 8%), AIRBAg (4/60, 7%), Aging Multidisciplinary Investigation (2/60, 3%) [[151](#)], BM3R (2/60, 3%) [[231](#)], FERMA (risk factors of the rural environment and allergic and respiratory disease; 1/60, 2%) [[232](#)], and Phytoneer (1/60, 2%) [[233](#)]. Of these, TRACTOR was the only cohort using exclusively MSA data [[222](#)].

Farming Exposure

A variety of exposure proxies were used to assess farming-related exposure. The most common proxy was a job title, which generally referred to whether the individual was a farmer (184/296, 62.2%). Other proxies included specific farming activities, such as dairy or crop farming (111/296, 37.5%), general pesticide exposure (yes or no; 62/296, 20.9%), and exposure to specific pesticide compounds (eg, glyphosate or paraquat) or pesticide classes (eg, insecticides; 51/296, 17.2%; Figure S10 in [Multimedia Appendix 1](#) and [Multimedia Appendix 2](#)). The number of farming activities studied ranged from just 1 [[226](#)] to 78 [[222](#)], with an average of 8 farming activities per study. Similarly, the number of pesticide compounds assessed ranged from 1 [[234](#)] to 943 [[235](#)], with an average of 42 pesticides per study. Only 1 study investigated the mixture effect of exposure to multiple pesticide combinations on human health [[236](#)]. Investigations into other chemical exposures were also limited, with only 2 papers each addressing silica exposure [[237,238](#)] and air pollution [[194,217](#)] (2/296, 0.7%). Notably, no studies examined exposure to per- and polyfluoroalkyl substances or micro- and nanoplastics. Research on the broader farming exposome was rare (5/296, 1.7%) and typically used farming activities as proxies [[152](#)].

Of 296 studies, few explored exposure to physical agents, with 5 studies (1.7%) focusing on radiation [[187,218,239-241](#)]. No studies investigated the effects of climate change on farmers' health. Exposure to biological agents was rarely studied as well, with just 3 (1%) out of 296 papers addressing mycotoxins [[241-243](#)]. Finally, only 3 studies (1%) examined psychological factors related to farming exposure [[244-246](#)].

Health Events

The most frequently studied health events were cancer (142/296, 48%), followed by mortality (44/296, 14.9%), injuries (38/296, 12.8%), workplace accidents (32/296, 10.8%), respiratory disorders (30/296, 10.1%), neurodegenerative diseases (28/296, 9.5%), and mental health issues (26/296, 8.8%; Figure S11 in [Multimedia Appendix 1](#) and [Multimedia Appendix 2](#)). Less frequently studied conditions included cardiovascular diseases (16/296, 5.4%), autoimmune disorders (11/296, 3.7%), musculoskeletal disorders (11/296, 3.7%), reproductive disorders (3/296, 1.0%), sleep disorders (1/296, 0.3%), and frailty (1/296, 0.3%). Notably, no studies explored the farming microbiome.

Among cancers, lung cancer was the most commonly investigated cancer (43/142, 30.3%), followed by prostate cancer (38/142, 26.8%), leukemia (37/142, 26.1%), colorectal cancer (35/142, 34.6%), multiple myeloma (35/142, 34.6%), non-Hodgkin lymphoma (35/142, 34.6%), bladder cancer (31/142, 21.8%), and brain cancer (31/142, 21.8%; Figure S12 in [Multimedia Appendix 1](#)). Respiratory disorders were primarily focused on asthma (15/30, 50%) and COPD (chronic obstructive pulmonary disease; 14/30, 47%). Parkinson disease was the most studied neurodegenerative condition (16/28, 57%), followed by multiple sclerosis (6/28, 21%). Fewer publications examined Alzheimer disease (2/28, 7%) and amyotrophic lateral sclerosis (2/28, 7%; Figure S13 in [Multimedia Appendix 1](#)). In the mental health field, suicide (12/26, 46%) and depression

(8/26, 31%) were the most investigated issues (Figure S14 in [Multimedia Appendix 1](#)).

Keyword Analysis

Overview

Following an initial extraction of 1259 authors' keywords, manual harmonization was performed. Duplicate keywords were removed through singular or plural standardization (130/1259, 10.3%) and synonym unification and grouping of cancer-related terms (553/1259, 43.9%), yielding a final set of 576 (45.8%) harmonized keywords, which were all used in subsequent analyses to prevent selection bias.

On average, each publication included 8.90 keywords (90% CI 0-17), although 35 (11.8%) out of 296 publications lacked any keywords, in line with the journal guidelines. Keyword analysis confirmed prior findings regarding farming exposure and health outcomes. It also provided deeper insights into emerging research hot spots, directions, and gaps.

Of the total 576 keywords, 301 (52.3%) appeared only once, while 68 (11.8%) were mentioned at least 10 times. More frequently used keywords included 39 that appeared at least 20 times (39/576, 6.8%) and 11 that featured in at least 50 publications (11/576, 1.9%). The 50 most frequently used keywords were mentioned in at least 17 (5.7%) out of 296 publications, while the top 10 appeared in at least 51 publications (17.2%; Table S7 in [Multimedia Appendix 1](#)). The most frequently cited keyword was "cancer" (150/296, 50.7%), followed by "mortality" (96/296, 32.4%), "pesticide" (88/296, 29.7%), "occupation" (82/296, 27.7%), "farmer" (77/296, 26.0%), "agriculture" (74/296, 25%), "exposure" (57/296, 19.3%), and "epidemiology" (57/296, 19.3%).

In terms of overall citations, "cancer" (5766/9379, 61.5%), "pesticide" (3569/9379, 38.1%), and "mortality" (3097/9379, 33%) were the most cited keywords. During the past decade, the frequency of the top 5 keywords has drastically increased (Figure S15 in [Multimedia Appendix 1](#)). Notably, keywords such as "cancer," "mortality," "occupation," "pesticide," "agriculture," and "farmer" have been consistently present in publications spanning at least 30 years (not necessarily consecutively; Figure S16 in [Multimedia Appendix 1](#)). In the

last decade, emerging keywords, such as "big data," "administrative health database," "dust," and "BMI," have gained prominence (Figure S17 in [Multimedia Appendix 1](#)).

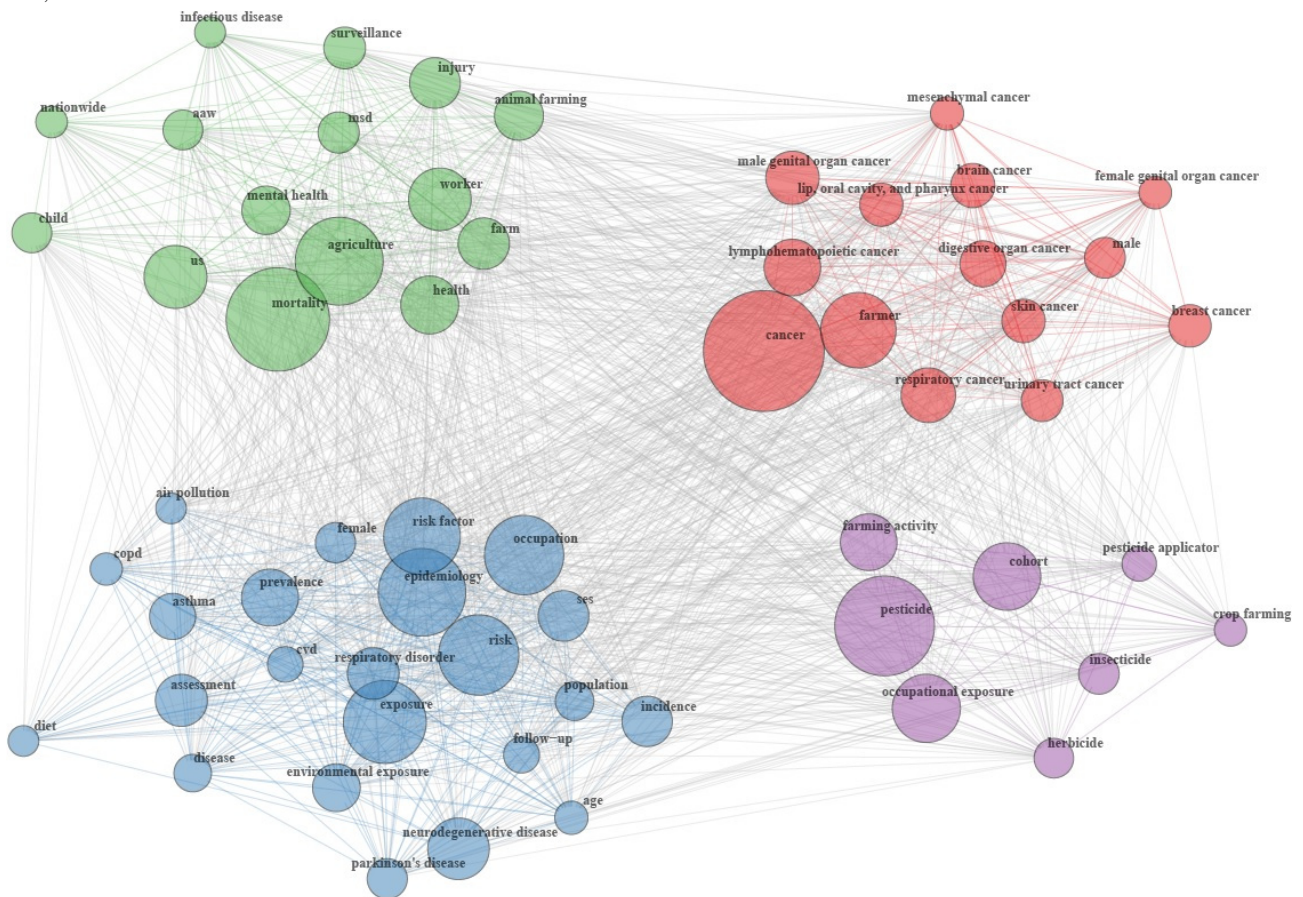
Keyword Co-Occurrence

A keyword co-occurrence network illustrating the frequency of keyword co-mentions in publications was constructed, thereby revealing relationships and conceptual connections (Figure 4). In this network, nodes or vertices represent keywords, with their sizes indicating frequency, while edges denote co-occurrences. The network's density and arrangement reveal topic interconnectivity, with larger vertices representing more frequently mentioned keywords. The network visualization helps identify clusters of related topics and highlights core research areas.

Using a community detection algorithm (spin-glass model with simulated annealing), 4 distinct clusters or communities of keywords were identified. Each cluster groups keywords that are often mentioned together, with stronger internal associations and weaker connections to keywords in other clusters.

The most frequently used keywords for each cluster were "cancer" (red cluster), "pesticide" (purple cluster), "mortality" (green cluster), and "exposure" (blue cluster). The red cluster highlights associations between various types of cancer, reflecting the fact that studies investigating cancer risks often examine multiple types of cancer. The green cluster links "mortality" with terms such as "mental health," "injury," and "animal farming," explained by the association between workplace accidents, mental health issues (eg, suicide), animal farming, and mortality. In the purple cluster, "pesticide" connects with "occupational exposure" and "farming activity," emphasizing that pesticide exposure is primarily studied in occupational settings across different types of farming. The blue cluster connects "exposure" to terms such as "neurodegenerative disease," "respiratory disorder," "cardiovascular disorder," "risk factor," "air pollution," "age," and "diet," indicating the study of various risk factors in relation to several health events. These clusters highlight current research hot spots that focus on 4 main interconnected themes: the associations between risk factors, pesticide exposure, farming activities, and a range of diseases.

Figure 4. Keyword co-occurrence network of the 296 articles published between 1975 and 2024. Each vertex or node represents a keyword, while edges represent the co-occurrence between keywords. Two keywords are connected when they co-occur in the same publication, and the size of each vertex indicates the frequency of a keyword: larger vertices represent more frequently mentioned keywords. Keywords with the same color (cluster) represent a research area. AAW: workplace accident; COPD: chronic obstructive pulmonary disease; CVD: cardiovascular disorder; MSD: musculoskeletal disorder; SES: socioeconomic status.



Thematic Mapping: Research Hot Spots

Figure 5 presents a thematic map that illustrates current research directions. Thematic mapping visualizes the relationship between research themes or topics, enabling the identification of directions, emerging areas, and gaps in the literature. The result is a strategic diagram that shows how themes relate to each other and their relevance within a specific field. The graph is divided into 4 quadrants, categorizing topics based on their relevance (x-axis, Callon centrality) and maturity (y-axis, Callon density) within the broader research landscape. Each circle represents a theme or topic (ie, a cluster of equivalent keywords), with the circle size corresponding to the frequency of the keywords associated with that theme.

The upper-right quadrant represents “hot topics,” which are both highly relevant and mature in the research landscape. Four key hot topics drive AHD-based public health research in farming populations. These include 1 topic focused on cancer research; another on respiratory disorders; and a third encompassing neurodegenerative diseases, workplace accidents, injuries, and mental health issues. The final hot topic involves large-scale studies in France and Europe using big data and insurance claims.

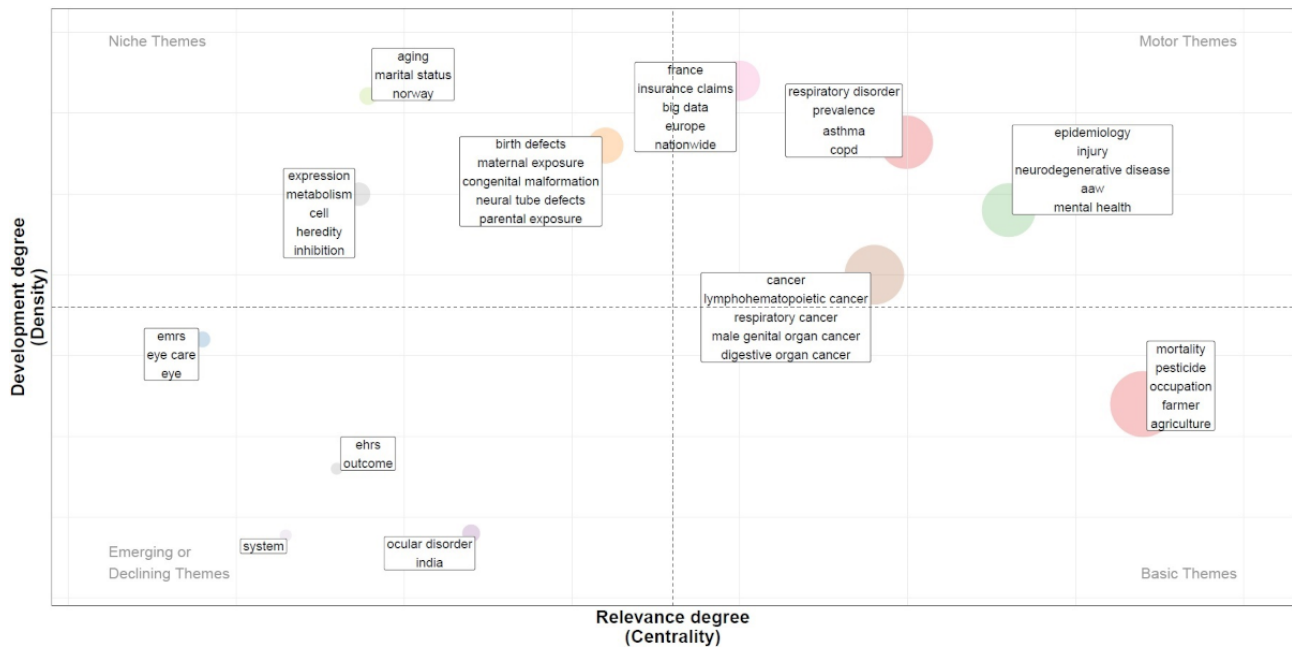
The lower-right quadrant contains “basic topics,” which are relevant but not yet mature in the research landscape. Only 1 such theme emerged: research related to pesticide exposure, mortality, and farming.

In the upper-left quadrant, “niche themes” refer to mature research topics that have not yet achieved full relevance. Three niche themes were identified: the first involves studies examining aging and research conducted in Norway; the second focuses on reproductive disorders and parental exposure, a theme poised to potentially evolve into a hot topic; and the final niche theme covers genetics and metabolism.

Finally, the lower-left quadrant contains “peripheral topics,” which represent either emerging or declining themes with low relevance and maturity. Four peripheral topics were observed, of which 2 (50%) were primarily centered on research on ocular disorders, 1 (25%) on the use of electronic health or medical records, and 1 (25%) on studies conducted in India.

This thematic map helps highlight both well-established and emerging areas of research, as well as gaps that may be ripe for future investigation.

Figure 5. Thematic mapping: research hot spots based on keywords from the 296 articles published between 1975 and 2024. The graph shows how themes relate to each other and their relevance within a specific field. This graph is divided into 4 quadrants, categorizing topics based on their relevance (x-axis and Callon centrality) and maturity (y-axis and Callon density) within the broader research landscape. Each circle represents a theme or topic (ie, a cluster of equivalent keywords), with the circle size corresponding to the frequency of the keywords associated with that theme. AAW: accident at work; COPD: chronic obstructive pulmonary disease; EHR: electronic health record; EMR: electronic medical record.



Bibliometric Farming Exposome

To identify research directions and gaps in the farming exposome literature, a bibliometric keyword co-occurrence analysis was conducted to explore the farming exposome by examining the co-occurrence between keywords associated with potential risk factors and specific health outcomes. This analysis was restricted to exposome-related and health event-related keywords. Of 576 keywords, 130 (22.6%) were related to the exposome, among which 93 (16.1%) were related to the specific external exposome (eg, pesticide), 19 (3.3%) to the general external exposome (eg, climate), and 18 (3.1%) to the internal

exposome (eg, oxidative stress). Furthermore, there were 70 (12.2%) health event-related keywords (eg, brain cancer).

The results of this analysis are provided in [Tables 5 and 6](#) and [Multimedia Appendix 3](#), with each cell representing the percentage of occurrences of an exposome-related keyword (eg, air pollution) in all publications mentioning a specific health event-related keyword (eg, Alzheimer disease). For example, a value of 33.3 indicates that an exposome-related keyword appeared in 33.3% of all publications mentioning a specified health event-related keyword. To facilitate interpretation and ease the reading of [Tables 5 and 6](#), exposome-related keywords were categorized into 19 groups (eg, chemical agent) and health event-related keywords into 20 groups.

Table 5. Co-occurrence between keywords related to internal exposomes and health event categories among the articles published between 1975 and 2024. Each cell refers to the number of times (%) a keyword related to an exposome category (eg, chemical agent) was mentioned among all publications in which a keyword related to a health event category (eg, cancer) appeared (N=296). Please note that the absolute value for each row is provided in parentheses with the row header and remains the same for all the parameters in that row.

Health event, n	Internal exposome (%)										
	Age	Sex	BMI	BP ^a	Heredity	Ethnicity	Hormone	Menopause	Metabolism	OS ^b	Inflammation
Cardiovascular disease (n=17)	5.88	11.8	5.88	5.88	0	5.88	0	0	0	0	0
Work-related disease (n=5)	0	20	0	0	0	0	0	0	0	0	0
Autoimmune disease (IBD ^c , RA ^d , vasculitis, and NR ^e ; n=9)	0	0	11.1	0	0	0	0	0	0	0	0
Cancer (n=150)	3.62	27.5	2.17	0	0.73	1.45	3.62	0.73	2.9	0.73	0
Dental health (n=2)	0	0	0	0	0	0	0	0	0	0	0
Ocular disorder (n=6)	16.7	0	0	0	0	0	0	0	0	0	0
Frailty (n=2)	50	0	50	0	0	0	0	0	0	0	0
Anemia (n=1)	100	100	0	0	0	0	0	0	100	0	0
Infectious disease (malaria, Lyme disease, tuberculosis, toxoplasmosis, and NR; n=14)	0	7.14	0	0	0	7.14	0	0	0	0	0
Injury (including workplace accident and disability; n=40)	4.26	8.51	2.13	0	0	2.13	0	2.13	0	0	0
Chronic kidney disease (n=3)	0	0	0	0	0	0	0	0	0	0	0
Mental health disorder (depression, suicide, and NR; n=25)	9.52	4.76	4.76	0	0	0	0	0	0	0	0
Metabolic disorder (diabetes, dysthyroidism, and NR; n=9)	11.1	0	11.1	11.1	0	11.1	0	0	0	0	0
Mortality (n=75)	4.17	17.7	0	0	1.04	0	1.04	2.08	0	0	0
Musculoskeletal disorder (arthritis, low-back pain, and NR; n=14)	7.14	14.3	7.14	0	0	0	0	0	0	0	0
Neurodegenerative disease (AD ^f , ALS ^g , MND ^h , MS ⁱ , PD ^j , and NR; n=33)	6.06	0	0	0	3.03	6.06	0	0	12.1	6.06	0
Sensory impairment (n=1)	0	0	0	0	0	0	0	0	0	0	0
Reproductive disorder (birth defects, infertility, spontaneous abortion, and NR; n=24)	0	28.6	14.3	0	0	7.14	7.14	0	7.14	0	0
Respiratory disorder (allergy, asthma, COPD ^k , pneumonia, sarcoidosis, and NR; n=39)	5.88	11.8	8.82	2.94	0	5.88	0	0	5.88	0	2.94
Skin disorder (dermatitis and NR; n=2)	0	50	0	0	0	0	0	0	0	0	0

^aBP: blood pressure.

^bOS: oxidative stress.

^cIBD: inflammatory bowel disease.

^dRA: rheumatoid arthritis.

^eNR: not reported.

^fAD: Alzheimer disease.

^gALS: amyotrophic lateral sclerosis.

^hMND: motor neuron disease

ⁱMS: multiple sclerosis.

^jPD: Parkinson disease.

^kCOPD: chronic obstructive pulmonary disease.

Distinct keyword exposome profiles were developed for each health event–related keyword (Figures S18-S43 in [Multimedia Appendix 1](#)), as illustrated in [Figure 6](#) for mental health disorders. Most exposome-related keywords associated with keywords related to mental health disorders pertained to the type of occupations as well as chemical, lifestyle, socioeconomic, and psychological factors. Cancer-related keywords were associated mostly with keywords related to the internal (sex) and specific external exposome (chemical agents, lifestyle, and type of occupations). Autoimmune disease–related keywords co-occurred mostly with external exposome–related keywords (chemical agents, lifestyle, type of occupations, and socioeconomic factors). Neurodegenerative disease–related keywords were associated mostly with keywords related to the specific external exposome (lifestyle, chemical agents, and type of occupations). Reproductive disorders co-occurred mostly

with internal (sex and BMI) and specific external exposome–related keywords (chemical agents and type of occupations). Keywords related to both musculoskeletal disorder and injury were associated with keywords from all 3 exposome components, in particular sex, type of occupations, lifestyle, biomechanical factors, chemical agents, and psychological factors. Infectious disease–related keywords co-occurred with specific external exposome–related keywords (biological agents and type of occupations), while respiratory disorder–related keywords were associated mostly with internal (sex) and specific external exposome–related keywords (lifestyle, chemical and biological agents, and type of occupations). Cardiovascular disorder–related keywords were associated with keywords from all 3 exposome components, in particular the sex, type of occupations, lifestyle, chemical agents, and socioeconomic and psychological factors.

Table 6. Co-occurrence between keywords related to specific external and general external exposomes and health event categories among the articles published between 1975 and 2024. Each cell refers to the number of times (%) a keyword related to an exposome category (eg, chemical agent) was mentioned among all publications in which a keyword related to a health event category (eg, cancer) appeared (N=296). Please note that the absolute value for each row is provided in parentheses with the row header and remains the same for all the parameters in that row.

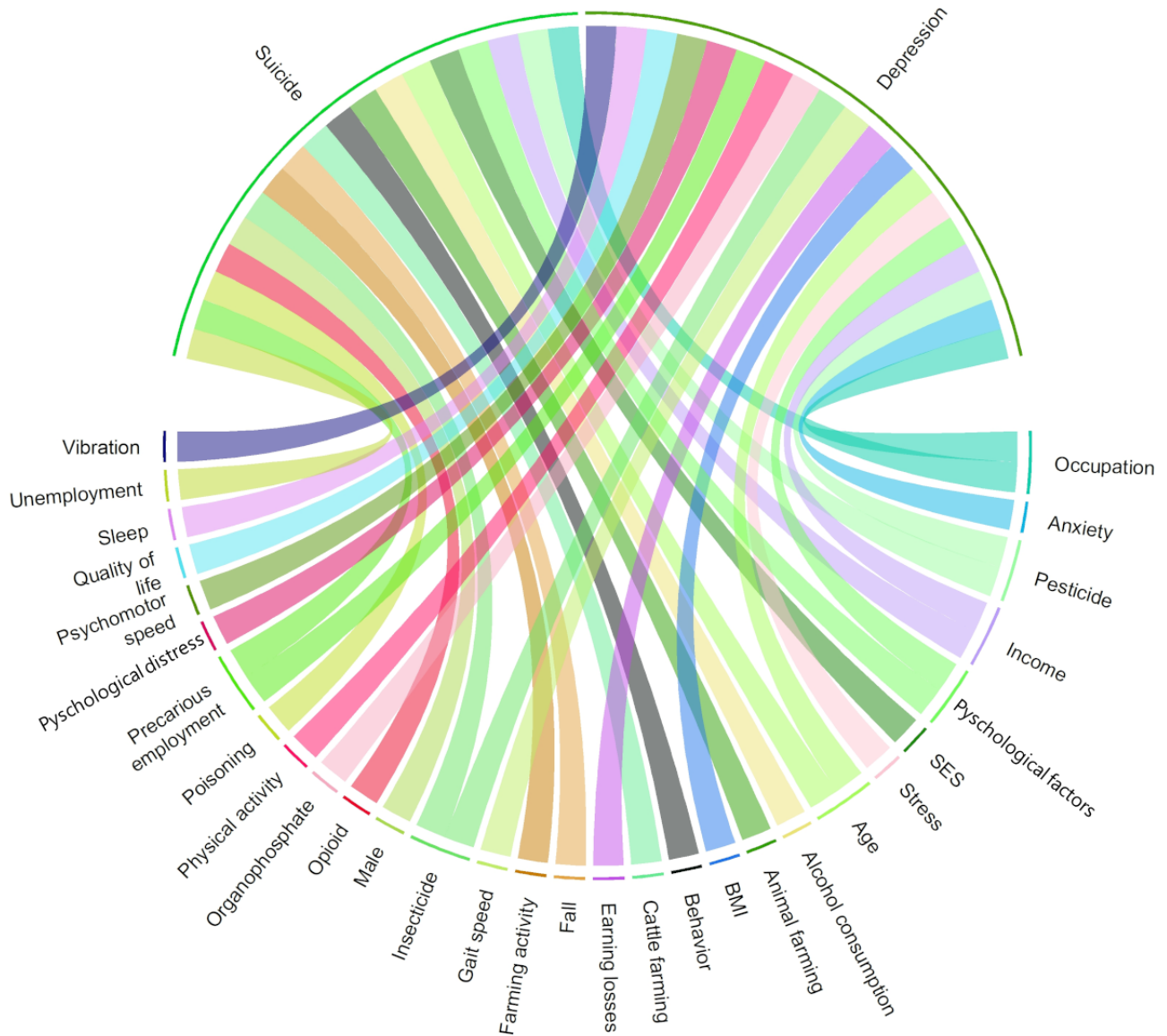
Health event, n	Specific external exposome (%)					General external exposome (%)		
	Lifestyle	CA ^a	BA ^b	BF ^c	Occupation	PA ^d	SF ^e	PF ^f
Cardiovascular disease (n=17)	11.8	29.4	0	11.8	23.5	0	29.4	17.6
Work-related disease (n=5)	0	20	20	20	40	0	20	0
Autoimmune disease (IBD ^g , RA ^h , vasculitis, and NR ⁱ ; n=9)	22.2	44.4	0	0	55.6	0	22.2	0
Cancer (n=150)	17.4	51.4	7.25	1.45	58.7	5.8	8.7	0
Dental health (n=2)	50	0	0	0	0	0	50	0
Ocular disorder (n=6)	16.7	16.7	0	0	0	0	0	16.7
Frailty (n=2)	100	50	0	0	0	0	0	50
Anemia (n=1)	0	0	0	0	0	0	0	0
Infectious disease (malaria, Lyme disease, tuberculosis, toxoplasmosis, and NR; n=14)	21.4	21.4	50	7.14	28.6	0	14.3	7.14
Injury (including workplace accident and disability; n=40)	12.8	14.9	2.13	14.9	23.4	2.13	4.26	6.38
Chronic kidney disease (n=3)	0	100	66.7	0	0	33.3	0	0
Mental health disorder (depression, suicide, and NR; n=25)	28.6	23.8	0	4.76	33.3	4.76	23.8	28.6
Metabolic disorder (diabetes, dysthyroidism, and NR; n=9)	33.3	44.4	0	11.1	44.4	11.1	0	22.2
Mortality (n=75)	17.7	55.2	8.33	4.17	54.2	5.21	17.7	2.08
Musculoskeletal disorder (arthritis, low-back pain, and NR; n=14)	14.3	21.4	7.14	14.3	42.9	7.14	0	21.4
Neurodegenerative disease (AD ^j , ALS ^k , MND ^l , MS ^m , PD ⁿ , and NR; n=33)	21.2	48.5	6.06	0	36.4	0	9.09	3.03
Sensory impairment (n=1)	0	0	0	0	100	0	0	0
Reproductive disorder (birth defects, infertility, spontaneous abortion, and NR; n=24)	7.14	35.7	7.14	0	64.3	0	7.14	0
Respiratory disorder (allergy, asthma, COPD ^o , pneumonia, sarcoidosis, and NR; n=39)	47.1	35.3	11.8	2.94	35.3	0	8.82	5.88
Skin disorder (dermatitis and NR; n=2)	50	50	0	0	50	0	0	0

^aCA: chemical agent.

^bBA: biological agent.

- ^cBF: biomechanical factor.
- ^dPA: physical agent.
- ^eSF: socioeconomic factor.
- ^fPF: psychological factor.
- ^gIBD: inflammatory bowel disease.
- ^hRA: rheumatoid arthritis.
- ⁱNR: not reported.
- ^jAD: Alzheimer disease.
- ^kALS: amyotrophic lateral sclerosis.
- ^lMND: motor neuron disease
- ^mMS: multiple sclerosis.
- ⁿPD: Parkinson disease.
- ^oCOPD: chronic obstructive pulmonary disease.

Figure 6. Chord diagram of keyword co-occurrence between potential risk factors and mental health disorder keywords from the 296 articles published between 1975 and 2024. Disease-related keywords are displayed on the top half of the chord diagram, while exposome-related keywords are displayed on the bottom half. Each chord or link indicates that an exposome-related keyword was mentioned with a disease-related keyword (co-occurrence) at least once in the same publication. The chord color differs from one exposome-related keyword to another. SES: socioeconomic status.



Discussion

Principal Findings

This review provides the first comprehensive and objective synthesis of research on the use of AHDs to address health issues in farming populations. It identifies major contributors, key publications, and existing research gaps while also suggesting future directions for leveraging AHDs to study health issues in farming populations. Overall, findings indicate that only a small part of the exposome and a limited range of health events have been examined within farming populations through the reuse of AHDs.

Current Directions in AHD Use for Public Health Research in Farming Populations

Research using AHDs in farming populations has been predominantly conducted in developed countries [150,225], with the United States [172,223,227]; France [207,222,224]; Canada [145,186,195]; and Scandinavian nations [182,189], including Denmark, Finland, Norway, and Sweden, leading the field. This dominance is linked to considerable funding from these regions and international collaborations. Scandinavian countries are particularly advanced in AHD use, offering databases that are highly complete, accessible, and well-integrated into public health research. AHDs from Denmark, Sweden, Canada, and France also provide comprehensive data on a patient's digital trajectory within their respective health systems [93,98,104,108,113,115,128,247]. France stands out further, with an AHD dedicated specifically to the entire farming population (MSA). This may explain the frequency of large-scale and long-term studies from these countries, some of which included >100,000 participants. However, many studies still had a regional focus, partly due to the use of AHDs by traditional epidemiological studies such as AGRICAN [207] and AHS [172], which rely on limited resources [47-50,53,54,57]. These studies often used AHDs to identify farming populations for inclusion or to supplement cohort data. The international AGRICOH consortium, initiated by the US National Cancer Institute and the International Agency for Research on Cancer, includes 11 (38%) of the 29 cohorts identified in this review [150,225]. However, several cohorts in AGRICOH were not identified, potentially due to lack of publications, language barriers, or limited use of AHDs. There were many publications associated with these well-established cohorts, for which many of the most prolific authors were working [172,182,189,207,222-227].

The most frequently used AHDs in farming population health research were disease registers, followed by electronic health or medical records and insurance claims. More than two-thirds of the studies used disease registers, in particular, cancer and mortality registers. This is not surprising because disease registers are created for clinical and research purposes with a continuous, exhaustive, and optimal digital collection of individual data regarding ≥ 1 health event in a geographically defined population [53,105]. The coding systems and the granularity of information related to health outcomes, populations, and exposure determinants varied widely across

studies. Most studies (291/296, 98.3%) used AHDs to collect sociodemographic and health event information [222].

There was no consensus on the best methods or proxies to assess farming exposure. A variety of exposure proxies and determinants were used across studies, with indirect methods being the most common. Many studies (237/296, 80.1%) dichotomized proxy, for example, classifying individuals as farmers or nonfarmers or as pesticide-exposed versus nonexposed. In nearly two-thirds of the included studies, job title (ie, "being a farmer") served as the primary exposure proxy. About one-third of studies took into account specific farming activities (eg, dairy farming and crop farming) to reflect the diversity of farming practices. This approach is a valuable proxy for agricultural exposure, offering a broader representation of the farming exposome, which involves multiple stressors beyond just pesticides [147,152,188,201,248,249]. Farming activity information was often derived from digitalized data, such as agricultural censuses or self-reported data from mandatory insurance enrollments [152]. Many studies (111/296, 37.5%) combined AHDs with self-reported data (eg, questionnaires) [172,207], which allowed for more comprehensive data collection but tended to restrict the scope to regional studies due to resource constraints. These studies typically yielded high-quality data, with more potential confounders considered compared to studies relying solely on AHDs. Most studies (68/111, 61.3%) using self-reported data focused on single exposures, mainly pesticides, with only 1 study addressing multiexposure to pesticides [236]. Biological monitoring and airborne chemical sampling were rarely conducted, likely due to practical and financial constraints and the short half-lives of most pesticides [250]. Dermal chemical monitoring has not been reported, even though it is the main exposure route for pesticides [251]. The high number of studies investigating exposure to pesticides may be explained by the fact that AHS focuses on pesticide applicators and their spouses [172] and because many pesticides have adverse health effects on humans, such as neurotoxicity or endocrine disruption [252-256]. Beyond pesticides, farmers face exposure to other chemicals [257], such as air pollution; micro- and nanoplastics [258-263]; biological agents (eg, endotoxins and zoonoses) [264-266]; physical agents (eg, UV radiation, noise, and vibrations) [187,267]; biomechanical factors (eg, repetitive movements, heavy load, and working posture) [198,268,269]; and psychosocial stressors [270-274], which have been less studied than pesticides. Despite these multiple exposures, the broader farming exposome remains understudied.

In addition to AHDs, some studies integrated other secondary data, such as climate data [187,275], air quality data [194,217], JEMs [212,276,277], or crop-exposure matrices [278]. JEMs provide exposure level estimates for various chemicals and stressors based on job categories [250,279]. Although JEMs can provide valuable exposure data, they often lack the specificity of individual-level data, making it difficult to account for task-specific risks, temporal variations, and the inclusion of specific worker subgroups such as female individuals [250,279-282]. The lack of a universal standard for JEMs further complicates their application, which may explain why many

studies still rely on self-reported data for more accurate exposure assessment despite the risk of recall bias [250,279-282].

The health outcomes studied were predominantly cancer [145,150,151,171-175,177-179,182,183,185,189,190,196,207,210-212,223], mortality [173,186,194,195,202,205], workplace injuries [198,208], respiratory disorders (eg, asthma and COPD) [151,180,181,213,224,226], neurodegenerative diseases (eg, Parkinson disease) [111,151,176,184,187,188,193,197,201,248,249,283], and mental health issues [151,244,273], which represent focal points within the research field. This is not surprising because many well-established cohorts centered on cancer research, in particular AGRICAN [207], Cancer in the Norwegian Agricultural Population [182], and NOCCA [189]. In addition, arsenic and inorganic arsenic compounds are classified as carcinogenic to humans by the International Agency for Research on Cancer [252,253,284], while malathion, glyphosate, diazinon, dichlorodiphenyltrichloroethane, and occupational exposures in spraying and application of nonarsenical insecticides are classified as probably carcinogenic to humans (group 2A), and several other pesticides are ranked as possibly carcinogenic to humans (group 2B), such as tetrachlorvinphos and parathion. Regarding mortality, it is often cancer and suicide mortalities that are investigated [186,195,202,244]. Furthermore, several pesticides are neurotoxic [252,255], but existing studies focused mainly on Parkinson disease and multiple sclerosis, with a paucity of data on Alzheimer disease and other neurodegenerative diseases [147,283]. In contrast, certain areas, such as cardiovascular diseases [151,194,203,227,285,286], autoimmune disorders [168,219,237], musculoskeletal disorders [192,204,245,287,288], reproductive disorders [242,289,290], sleep disorders [191], aging-related conditions [151,291], hearing impairment [267,292], and the microbiome [8,293-298], remain underexplored, despite their potential relevance to farming populations.

Challenges of Reusing AHD for Public Health Research in Farming Populations

Each AHD presents unique advantages and limitations. For example, large sample sizes and a large number of available health events are strengths, while generalizability and the absence of key confounders are challenges [64,93,95,105,115]. Access to AHDs is frequently restricted by a variety of factors, including governance and technical barriers, such as language, data structure, interoperability, and coding systems. Additional challenges stem from the type of AHD (eg, insurance claims or cancer registers), inadequate documentation (eg, absence of a data dictionary), limited accessibility due to costs or conditions, and jurisdictional and legal constraints [30,62,64,81,113,115]. Identifying the optimal AHD for a given research question is also complex, especially when considering the heterogeneity in coding systems and country-specific data structures. In countries such as Scandinavia, Canada, and France, individual identifiers facilitate data linkage across multiple AHDs, enhancing research opportunities [93,104,113,115,128,247]. However, many AHDs are not ready for research and require significant processing, cleaning, and understanding before they can be analyzed [93,105,113,222,299]. Another major challenge is the long lag between data access, analysis, and research

publication. On average, studies used data that were 7 years old at the time of publication, largely due to delays in data access, administrative approvals, and the need to prepare complex datasets for analysis [223,300]. For instance, the TRACTOR project took 2 years to clean and prepare its dataset for research use [222]. These delays are compounded by the time required to conduct statistical analyses and prepare manuscripts for publication, as well as review and publication times (delay from submission to acceptance and from acceptance to publication). Another limitation of AHDs is the lack of detailed exposure data. AHDs rarely include exposure information due to their administrative focus, requiring researchers to supplement with additional data sources, such as JEMs or self-reported data. When exposure information is recorded in AHDs, it is often too generalized, typically only reflecting broad job classifications, such as farming, without specifying detailed activities or stressors. There are some exceptions, such as MSA data, which capture a wider range of specific farming activities (eg, dairy farming) [222]. However, exposure to specific stressors (eg, chemical compounds) is largely absent from AHDs.

The reference populations used in farming studies vary, which precludes direct comparisons and limits the generalizability of the findings. For example, AGRICAN used the general population as a reference [207], while TRACTOR used a farming population [152,168]. Furthermore, studies differ in their focus on specific farming populations, such as the entire agricultural workforce [207], farm managers [152], or pesticide applicators [172], which may lead to distinct exposure profiles that influence health outcomes because these farming populations have different socioeconomic status, experiences, and behaviors. Hence, to avoid or lessen bias, some studies focused on specific farming populations [152,172]. Moreover, the scope of farming populations included in studies is often limited, omitting subgroups, such as farm families, nearby residents, or consumers exposed to agricultural products, which limits the broader application of the findings. In addition, farming practices can vary significantly between countries and studies, and there is no international standardized classification for farming activities. In some cases, farming categories are derived from legal or administrative sources, as seen in the MSA data [152,168,283]. This lack of standardization limits the comparability and generalizability of findings across studies. In addition, the generalizability of the findings to other countries when using AHDs may be limited because of the differences in health care systems [93].

There are several well-known limitations of AHDs that complicate the investigation of health outcomes [301]. Health events captured in AHDs are typically limited to those requiring medical attention, which may not reflect the true incidence of diseases. In addition, the level of detail varies across diseases, even within the same AHD [92]. Although diagnostic accuracy is generally higher in disease registers, these are often geographically limited and cover only a subset of health conditions. For example, in France, cancer registers only cover 23 (24%) out of 96 administrative regions [155]. In addition, certain conditions, such as depression, are not covered by any registers. Identifying health outcomes in AHDs often requires complex algorithms that combine data from multiple sources,

such as drug reimbursements, disease declarations, or medical procedures [104,105,128,152,302-304]. In addition, inconsistencies in case definitions and algorithms across studies and countries hinder the ability to compare and pool risk estimates [104,302,304]. Some AHDs also lack critical clinical information, such as laboratory results and genetic data [115,128,305], and the recorded diagnosis or treatment date may not correspond to the actual onset of the disease. Furthermore, diagnosis codes are not always indicative of a confirmed diagnosis.

Emerging Opportunities and Research Needs

While AHDs are well-used in certain countries, there are underexplored opportunities in regions such as the United Kingdom, where AHDs exist but are underused for research [113]. For low- and middle-income countries, the development and access to AHDs remain limited, and international support is needed to expand this research infrastructure. As already reported by Habib et al [306], there is a notable lack of sex-specific data, even though occupational exposures and health outcomes can vary significantly between sexes due to genetic, physiological, psychological, and behavioral factors [307-311]. Future research should address these disparities to provide a more comprehensive understanding of health risks in farming communities. Although there are inherent delays in using AHDs due to the time required for data generation, consolidation, and access, we advocate for the continued publication of studies, even those using older data. Historical data remain vital for better understanding long-term health patterns, particularly for diseases such as cancer, where tumor initiation can span decades [312,313]. Editors should encourage the publication of studies using older datasets, especially when addressing long-term health outcomes (eg, cancer and neurodegenerative diseases) or when recent data are not available [312].

None of the analyzed AHDs fully adhere to the FAIR principles, possibly because most were developed before the establishment of these principles in 2016 [228-230]. Moreover, the assessment of FAIR compliance of AHDs relied solely on information presented in publications, which may not provide a comprehensive evaluation. Nevertheless, there is a critical need to advocate for the integration of FAIR principles within AHDs to enhance public health research [228-230]. Currently, the landscape is favorable for data reuse, particularly with initiatives such as the forthcoming European Health Data Space [314-316]. Data reuse extends beyond mere access; it encompasses data discovery, a fundamental aspect of FAIR principles that involves recognizing the existence of databases [228-230]. To facilitate this, the creation of data catalogs is essential [228]. Numerous data repositories, such as Re3Data [317], Zenodo [318], CANUE [319], Figshare [320], “Epidémiologie – France” [321], data.gouv [322], Dataverse [320], or Data Europa [320], already exist. In addition, specialized multidisciplinary open-access and peer-reviewed journals such as *Scientific Data* and *Data in Brief* publish datasets [318]. A dataset search can also be conducted using the Google (Google LLC) platform [323]. However, the documentation and access conditions for datasets can highly vary across inventories, complicating the selection process for researchers. The absence of indicators or scores for data reusability further hampers efforts to identify the most suitable

datasets for specific research questions [45,69]. To our knowledge, no comprehensive catalog of AHDs currently exists to date. A web-based inventory of AHDs, modeled after existing resources, such as OccupationalCohorts.net [324], OccupationalExposureTools.net [325], and Toxicological and Exposure Database Inventory [10], could greatly enhance research endeavors. The motivation for analyzing AHDs often stems from the data they contain. Consequently, as data availability increases, researchers will be better positioned to formulate research questions and engage in a parallel process of “datagraphy” or “datagraphic search” akin to traditional bibliographic research. The objective of datagraphy would be to determine which datasets are best suited for addressing specific research questions, highlighting the need for accessible catalogs to support this goal.

There is also an opportunity to integrate other secondary data, such as person-generated data (eg, mobile health, social media, digital footprints, and wearable sensors); contextual data (eg, climate and air quality); and smart agriculture data [2,83,84,326-328]. These datasets, largely untapped in farming population research, could provide new insights into health outcomes and environmental exposures [101].

Nationwide studies using big data were a hot spot. AI, such as machine learning, is particularly useful for analyzing big data and holds substantial promise for future research [329-331], particularly for identifying predictors of health outcomes in farming populations [332]. To date, AI has been underused, with only 2 studies using it, 1 identifying occupational injuries in agriculture [208] and 1 reviewing the development of chronic kidney disease risk prediction models [209]. Incorporating AI, along with cohort enrichment and interdisciplinary expert interpretation, could open new avenues for research.

Many studies continue to examine agriculture as a broad category, highlighting the need for more detailed investigations into specific farming activities and tasks [147,152,188,201,248,249]. Our analysis reveals major research gaps in understanding environmental and occupational exposures among farming populations, particularly with regard to emerging concerns such as per- and polyfluoroalkyl substances, biological agents, micro- and nanoplastics, and the impact of climate change. Climate change is a critical issue for agriculture, as it may drive shifts in pests, diseases, and farming practices [274,333-339]. Parental exposure appears to be a theme that will soon become a hot topic. Furthermore, research is needed to explore the farming exposome, particularly focusing on the “mixture effects” of multiple simultaneous exposures [340,341]. Omics data, which have not been used in farming population studies to date, represent a promising avenue for future research because genetics and metabolism were found to be a niche theme. Omics data refer to the large-scale datasets generated from various omics technologies that analyze biological molecules (eg, genomics, transcriptomics, proteomics, and metabolomics), which provide comprehensive insights into different biological layers and processes [11,342,343].

To enhance the characterization of farming exposome research using keyword analysis, there is a pressing need for standardized keyword reporting. We advocate for the development of a

standardized approach to reporting keywords in scientific journals, including defining a minimum set of information (eg, study type, health outcome, population studied, data sources, and positive, negative, or null associations) and adopting a list of standardized terms. Although challenging, this approach would improve literature searches, make data more comparable and FAIR [228-230], and lead to more efficient, frugal (less time and energy spent to identify relevant information), and accurate synthesis of the scientific literature, such as in reviews and bibliometric analyses.

The prominence of topics such as cancer, neurodegenerative diseases, mortality, injuries, and mental health issues underscores the need for targeted prevention strategies. The thematic map analysis indicates that reproductive disorders (eg, birth defects, endometriosis, and infertility) are on the verge of becoming a central research focus. Emerging and understudied health conditions, including ocular disorders, autoimmune diseases (eg, inflammatory bowel disease and rheumatoid arthritis), sleep disorders (eg, sleep apnea), cardiovascular diseases, and musculoskeletal conditions (eg, low-back pain), warrant increased attention and further research. Aging-related health issues, such as frailty, also represent promising avenues for future research, particularly given the growing aging population and associated health care challenges [24].

Limitations

The findings of this review should be considered in view of their limitations. Because of time and resource constraints, a single screening approach was used, with only 1 author (PP) conducting the review and bibliometric analysis. While single screening is an efficient use of time and resources, there is a higher risk of missing relevant studies than when using dual screening [344,345]. However, when completed by an experienced reviewer familiar with the research topic, the proportion of missed studies is limited and estimated to be around 3% [344]. Therefore, we cannot exclude the possibility that some studies may have been missed. Nevertheless, we are confident that none of these methodological limitations would change the overall conclusions of this work. Our restriction on articles published in English and French may have inadvertently excluded potentially relevant publications. We cannot exclude the possibility that publications using AHDs for addressing health issues in farming populations may have been missed if there was no mention of AHD in the publications' titles and abstracts. However, it is important to mention that our search strategy was similar to recent reviews that specifically examined the use of AHDs for population-based research [93,96,103,104]. We further broadened our search by including synonyms to improve the comprehensiveness of our literature search. Some details and specificities on the AHDs and other digital data used may be limited because only information reported in each study was used. Shortcomings inherent to bibliometric analysis cannot be excluded. Some authors may have duplicate names, and namesakes could exist. This limitation could not be prevented as a unique author identifier (eg, Open Researcher and

Contributor ID number) was not available. Self-citation could not be identified.

While our keyword analysis helped map the farming exposome in AHD-based public health research, this profile is incomplete and potentially biased. Because our review focused on AHD-based studies, we likely missed relevant epidemiological studies not using AHDs, leaving gaps in our understanding of the complete farming exposome across public health. In addition, the variability in keyword reporting practices across journals introduced bias into our keyword analysis. Some journals limit the number of keywords, and the lack of standardized keyword ontologies adds further variability. To mitigate this bias, we manually harmonized the keywords (eg, the use of 1 unique term for a given entity). While this approach is time-consuming, it allows for a more accurate analysis. For instance, if this approach was not performed, the same entity could be designated by various terms that would have been considered separate entities or terms, potentially resulting in underestimates (eg, in the number of publications). Despite these challenges, the findings from our scoping review were consistent with the keyword analysis.

Notwithstanding the aforementioned limitations, most of which are inherent to all scoping reviews and bibliometric analyses [93,96,103,104,158-160], we are confident that our findings can provide a comprehensive picture of what has been published until now (the current state of research and general directions) regarding the use of AHDs for addressing health issues in farming populations. This study may lay the groundwork for researchers to quickly identify research priorities and emerging research directions investigating health issues in farming populations using AHDs.

Conclusions

Technological advancements have greatly increased the volume of research data available, positioning AHDs as critical resources for population-based public health studies [41]. Our review underscores the broad public health implications of AHDs, providing actionable insights for researchers, physicians, and policy makers (Textbox 3). Addressing the identified research gaps is crucial to comprehensively understanding health risks in farming populations.

The insights derived from AHDs can inform meaningful recommendations for policy makers and guide future research directions, ultimately aiding health services and health policy development. Our findings underscore the necessity of comprehensive, interdisciplinary approaches to better understand and mitigate the health risks encountered by farming populations. Such efforts will improve data comparability and research quality while also supporting the formulation of targeted prevention strategies. This, in turn, will enhance health outcomes for farming populations and promote the sustainability of agriculture in an increasingly dynamic environment. The findings from this review offer insights that are not only relevant to farming populations but also potentially generalizable to other populations.

Textbox 3. Take-home messages.

Farming population

Research focusing on low- and middle-income countries, as well as on underrepresented subgroups within farming communities (eg, women, children, and contingent workers), remains insufficiently developed. These areas warrant further investigation to ensure more comprehensive insights.

Administrative health database (AHD) use

The use of AHDs in public health research among farming populations is expanding, offering major potential to enhance epidemiological studies and inform public health decisions. Promoting AHD-based research alongside the integration of other secondary data and artificial intelligence approaches represents a promising direction for future exploration. There is also a need to promote findable, accessible, interoperable, and reusable principles. Creating an AHD catalog or inventory could be a solution that would allow researchers to conduct a “datagraphic search” akin to traditional bibliographic research.

Farming exposure

Published studies on farming-related exposures often rely on broad proxies, such as job titles, neglecting to capture the nuances of specific agricultural tasks. While pesticide exposure remains a predominant research focus, emerging concerns, such as per- and polyfluoroalkyl substances, biological agents, micro or nanoplastics, and the effects of climate change, require urgent attention. The farming exposome remains underexplored despite its potential to uncover important associations between risk factors and a diverse range of health outcomes.

Health outcomes

Cancer, respiratory diseases, neurodegenerative disorders, and mental health issues are among the most frequently studied health outcomes in farming populations. However, significant gaps exist in understanding other critical conditions, such as cardiovascular diseases, reproductive disorders, ocular conditions, autoimmune diseases, musculoskeletal disorders, age-related health issues, and microbiome impacts. Addressing these overlooked areas is essential for a more complete understanding of the health risks faced by farming communities.

Acknowledgments

This work was partly funded by MIAI@Grenoble Alpes (Multidisciplinary Institute in Artificial Intelligence at Grenoble Alpes; grant ANR-19-P3IA-0003, 2019) and the French National Research Agency within the framework of the Investissements d’avenir program (grants ANR - 10 - AIRT - 05, 2012; ANR - 15 - IDEX - 02, 2016). The funding sources had no role in the study design and conduct of the study; collection, management, analysis, and interpretation of data; writing of the report; preparation, review, or approval of the paper; or decision to submit the paper for publication. The authors were not precluded from accessing data in the study, and all authors took final responsibility for the decision to submit for publication.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request and also in the multimedia appendices.

Authors' Contributions

PP was involved in the conceptualization, methodology, software selection, validation, formal analysis, investigation, data curation, writing the original draft and reviewing and editing, visualization, project administration, and funding acquisition of the study. NV was involved in the conceptualization, reviewing and editing, resource collection, and funding acquisition of the study. All authors meet all 4 criteria for authorship in the International Committee of Medical Journal Editors recommendations. All authors affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any potential discrepancies from the study as originally planned (and, if relevant, registered) have been explained.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary materials.

[\[DOCX File, 8778 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Publications included and analyzed.

[\[XLSX File \(Microsoft Excel File\), 80 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Co-occurrence between exposome-related keywords and health event-related keywords.

[[XLSX File \(Microsoft Excel File\), 39 KB-Multimedia Appendix 3](#)]

Multimedia Appendix 4

PRISMA-ScR checklist.

[[PDF File \(Adobe PDF File\), 186 KB-Multimedia Appendix 4](#)]

Multimedia Appendix 5

BIBLIO checklist for reporting the bibliometric reviews of the biomedical literature.

[[PDF File \(Adobe PDF File\), 270 KB-Multimedia Appendix 5](#)]

References

1. Azari R, Borisch B. What is public health? A scoping review. *Arch Public Health*. May 10, 2023;81(1):86. [[FREE Full text](#)] [doi: [10.1186/s13690-023-01091-6](https://doi.org/10.1186/s13690-023-01091-6)] [Medline: [37165370](#)]
2. Velmovitsky PE, Kirolos M, Alencar P, Leatherdale S, Cowan D, Morita PP. Leveraging mHealth technologies for public health. *JMIR Public Health Surveill*. Sep 12, 2024;10:e49719. [[FREE Full text](#)] [doi: [10.2196/49719](https://doi.org/10.2196/49719)] [Medline: [39265164](#)]
3. Barouki R, Lévi Y. Rapport 24-02. L'exposome, une contribution majeure pour la prévention. *Bull Acad Natl Med*. May 2024;208(5):523-535. [doi: [10.1016/j.banm.2024.03.004](https://doi.org/10.1016/j.banm.2024.03.004)]
4. Barouki R, Audouze K, Becker C, Blaha L, Coumoul X, Karakitsios S, et al. The exposome and toxicology: a win-win collaboration. *Toxicol Sci*. Feb 28, 2022;186(1):1-11. [[FREE Full text](#)] [doi: [10.1093/toxsci/kfab149](https://doi.org/10.1093/toxsci/kfab149)] [Medline: [34878125](#)]
5. Bennett B, Workman T, Smith MN, Griffith WC, Thompson B, Faustman EM. Characterizing the neurodevelopmental pesticide exposome in a children's agricultural cohort. *Int J Environ Res Public Health*. Feb 25, 2020;17(5):1479. [[FREE Full text](#)] [doi: [10.3390/ijerph17051479](https://doi.org/10.3390/ijerph17051479)] [Medline: [32106530](#)]
6. Goodrich JA, Wang H, Jia Q, Stratakis N, Zhao Y, Maitre L, et al. Integrating Multi-Omics with environmental data for precision health: a novel analytic framework and case study on prenatal mercury induced childhood fatty liver disease. *Environ Int*. Aug 2024;190:108930. [[FREE Full text](#)] [doi: [10.1016/j.envint.2024.108930](https://doi.org/10.1016/j.envint.2024.108930)] [Medline: [39128376](#)]
7. Haddad N, Andrianou XD, Makris KC. A scoping review on the characteristics of human exposome studies. *Curr Pollution Rep*. Nov 28, 2019;5(4):378-393. [doi: [10.1007/S40726-019-00130-7](https://doi.org/10.1007/S40726-019-00130-7)]
8. Merra G, Gualtieri P, La Placa G, Frank G, Della Morte D, De Lorenzo A, et al. The relationship between exposome and microbiome. *Microorganisms*. Jul 08, 2024;12(7):1386. [[FREE Full text](#)] [doi: [10.3390/microorganisms12071386](https://doi.org/10.3390/microorganisms12071386)] [Medline: [39065154](#)]
9. Pandics T, Major D, Fazekas-Pongor V, Szarvas Z, Peterfi A, Mukli P, et al. Exposome and unhealthy aging: environmental drivers from air pollution to occupational exposures. *Geroscience*. Dec 09, 2023;45(6):3381-3408. [[FREE Full text](#)] [doi: [10.1007/s11357-023-00913-3](https://doi.org/10.1007/s11357-023-00913-3)] [Medline: [37688657](#)]
10. Petit P. Toxicological and exposure database inventory: a review. *Int J Hyg Environ Health*. Sep 2022;246:114055. [doi: [10.1016/j.ijheh.2022.114055](https://doi.org/10.1016/j.ijheh.2022.114055)] [Medline: [36279787](#)]
11. Price EJ, Vitale CM, Miller GW, David A, Barouki R, Audouze K, et al. Merging the exposome into an integrated framework for "omics" sciences. *iScience*. Mar 18, 2022;25(3):103976. [[FREE Full text](#)] [doi: [10.1016/j.isci.2022.103976](https://doi.org/10.1016/j.isci.2022.103976)] [Medline: [35310334](#)]
12. Santos S, Maitre L, Warembourg C, Agier L, Richiardi L, Basagaña X, et al. Applying the exposome concept in birth cohort research: a review of statistical approaches. *Eur J Epidemiol*. Mar 27, 2020;35(3):193-204. [[FREE Full text](#)] [doi: [10.1007/s10654-020-00625-4](https://doi.org/10.1007/s10654-020-00625-4)] [Medline: [32221742](#)]
13. Siroux V, Agier L, Slama R. The exposome concept: a challenge and a potential driver for environmental health research. *Eur Respir Rev*. Jun 31, 2016;25(140):124-129. [[FREE Full text](#)] [doi: [10.1183/16000617.0034-2016](https://doi.org/10.1183/16000617.0034-2016)] [Medline: [27246588](#)]
14. Tagliaferro S, Maio S, Pirona F, Stanisci I, Sarno G, Silvi P, EarlyFOOD, HEALS EXHES, et al. Assessing external exposome by implementing an environmental data management system using open data. *Sci Rep*. Jul 26, 2024;14(1):17142. [[FREE Full text](#)] [doi: [10.1038/s41598-024-62924-0](https://doi.org/10.1038/s41598-024-62924-0)] [Medline: [39060268](#)]
15. Vineis P, Barouki R. The exposome as the science of social-to-biological transitions. *Environ Int*. Jul 2022;165:107312. [[FREE Full text](#)] [doi: [10.1016/j.envint.2022.107312](https://doi.org/10.1016/j.envint.2022.107312)] [Medline: [35635963](#)]
16. Wei X, Huang Z, Jiang L, Li Y, Zhang X, Leng Y, et al. Charting the landscape of the environmental exposome. *Imeta*. Dec 02, 2022;1(4):e50. [doi: [10.1002/imt2.50](https://doi.org/10.1002/imt2.50)] [Medline: [38867899](#)]
17. Wild CP. The exposome: from concept to utility. *Int J Epidemiol*. Feb 2012;41(1):24-32. [doi: [10.1093/ije/dyr236](https://doi.org/10.1093/ije/dyr236)] [Medline: [22296988](#)]
18. Zhang P, Carlsten C, Chaleckis R, Hanhineva K, Huang M, Isobe T, et al. Defining the scope of exposome studies and research needs from a multidisciplinary perspective. *Environ Sci Technol Lett*. Oct 12, 2021;8(10):839-852. [[FREE Full text](#)] [doi: [10.1021/acs.estlett.1c00648](https://doi.org/10.1021/acs.estlett.1c00648)] [Medline: [34660833](#)]

19. Ibanez A, Melloni L, Świeboda P, Hynes W, Ikiz B, Ayadi R, et al. Neuroecological links of the exposome and one health. *Neuron*. Jun 19, 2024;112(12):1905-1910. [FREE Full text] [doi: [10.1016/j.neuron.2024.04.016](https://doi.org/10.1016/j.neuron.2024.04.016)] [Medline: [38723637](https://pubmed.ncbi.nlm.nih.gov/38723637/)]
20. Liang KH, Colombijn JM, Verhaar MC, Ghannoum M, Timmermans EJ, Vernooij RW. The general external exposome and the development or progression of chronic kidney disease: a systematic review and meta-analyses. *Environ Pollut*. Oct 01, 2024;358:124509. [FREE Full text] [doi: [10.1016/j.envpol.2024.124509](https://doi.org/10.1016/j.envpol.2024.124509)] [Medline: [38968981](https://pubmed.ncbi.nlm.nih.gov/38968981/)]
21. Canfell OJ, Littlewood R, Burton-Jones A, Sullivan C. Digital health and precision prevention: shifting from disease-centred care to consumer-centred health. *Aust Health Rev*. Jun 10, 2022;46(3):279-283. [doi: [10.1071/AH21063](https://doi.org/10.1071/AH21063)] [Medline: [34882538](https://pubmed.ncbi.nlm.nih.gov/34882538/)]
22. Mess F, Blaschke S, Schick TS, Friedrich J. Precision prevention in worksite health—a scoping review on research trends and gaps. *PLoS One*. Jun 10, 2024;19(6):e0304951. [FREE Full text] [doi: [10.1371/journal.pone.0304951](https://doi.org/10.1371/journal.pone.0304951)] [Medline: [38857277](https://pubmed.ncbi.nlm.nih.gov/38857277/)]
23. Wang Y, Zhu M, Ma H, Shen H. Polygenic risk scores: the future of cancer risk prediction, screening, and precision prevention. *Med Rev (2021)*. Dec 2021;1(2):129-149. [FREE Full text] [doi: [10.1515/mr-2021-0025](https://doi.org/10.1515/mr-2021-0025)] [Medline: [37724297](https://pubmed.ncbi.nlm.nih.gov/37724297/)]
24. Kalia V, Belsky DW, Baccarelli AA, Miller GW. An exposomic framework to uncover environmental drivers of aging. *Exposome*. 2022;2(1):osac002. [FREE Full text] [doi: [10.1093/exposome/osac002](https://doi.org/10.1093/exposome/osac002)] [Medline: [35295547](https://pubmed.ncbi.nlm.nih.gov/35295547/)]
25. Chung MK, House JS, Akhtari FS, Makris KC, Langston MA, Islam KT, et al. Members of the Exposomics Consortium. Decoding the exposome: data science methodologies and implications in exposome-wide association studies (ExWASs). *Exposome*. 2024;4(1):osae001. [FREE Full text] [doi: [10.1093/exposome/osae001](https://doi.org/10.1093/exposome/osae001)] [Medline: [38344436](https://pubmed.ncbi.nlm.nih.gov/38344436/)]
26. Andreu-Perez J, Poon CC, Merrifield RD, Wong ST, Yang GZ. Big data for health. *IEEE J Biomed Health Inform*. Jul 2015;19(4):1193-1208. [FREE Full text] [doi: [10.1109/JBHI.2015.2450362](https://doi.org/10.1109/JBHI.2015.2450362)] [Medline: [26173222](https://pubmed.ncbi.nlm.nih.gov/26173222/)]
27. Li C, Huang M. Environmental sustainability in the age of big data: opportunities and challenges for business and industry. *Environ Sci Pollut Res Int*. Dec 03, 2023;30(56):119001-119015. [doi: [10.1007/s11356-023-30301-5](https://doi.org/10.1007/s11356-023-30301-5)] [Medline: [37919492](https://pubmed.ncbi.nlm.nih.gov/37919492/)]
28. Pastorino R, De Vito C, Migliara G, Glocker K, Binenbaum I, Ricciardi W, et al. Benefits and challenges of Big Data in healthcare: an overview of the European initiatives. *Eur J Public Health*. Oct 01, 2019;29(Supplement_3):23-27. [FREE Full text] [doi: [10.1093/eurpub/ckz168](https://doi.org/10.1093/eurpub/ckz168)] [Medline: [31738444](https://pubmed.ncbi.nlm.nih.gov/31738444/)]
29. Flahault A, Geissbuhler A, Guessous I, Guérin P, Bolon I, Salathé M, et al. Precision global health in the digital age. *Swiss Med Wkly*. Apr 07, 2017;147(1314):w14423. [FREE Full text] [doi: [10.4414/sm.w.2017.14423](https://doi.org/10.4414/sm.w.2017.14423)] [Medline: [28421566](https://pubmed.ncbi.nlm.nih.gov/28421566/)]
30. Manaouil C, Chamot S, Petit P. Le médecin confronté à l'IA (Intelligence artificielle): Éthique et responsabilité. *Méd Droit*. Jun 2024;2024(186):50-66. [doi: [10.1016/j.meddro.2024.02.001](https://doi.org/10.1016/j.meddro.2024.02.001)]
31. Samuel G, Lucassen AM. The environmental sustainability of data-driven health research: a scoping review. *Digit Health*. Jul 05, 2022;8:20552076221111297. [FREE Full text] [doi: [10.1177/20552076221111297](https://doi.org/10.1177/20552076221111297)] [Medline: [35847526](https://pubmed.ncbi.nlm.nih.gov/35847526/)]
32. Adam NR, Wieder R, Ghosh D. Data science, learning, and applications to biomedical and health sciences. *Ann N Y Acad Sci*. Jan 25, 2017;1387(1):5-11. [doi: [10.1111/nyas.13309](https://doi.org/10.1111/nyas.13309)] [Medline: [28122121](https://pubmed.ncbi.nlm.nih.gov/28122121/)]
33. Iyamu I, Xu AX, Gómez-Ramírez O, Ablona A, Chang H, Mckee G, et al. Defining digital public health and the role of digitization, digitalization, and digital transformation: scoping review. *JMIR Public Health Surveill*. Nov 26, 2021;7(11):e30399. [FREE Full text] [doi: [10.2196/30399](https://doi.org/10.2196/30399)] [Medline: [34842555](https://pubmed.ncbi.nlm.nih.gov/34842555/)]
34. Mooney SJ, Pejaver V. Big data in public health: terminology, machine learning, and privacy. *Annu Rev Public Health*. Apr 01, 2018;39:95-112. [FREE Full text] [doi: [10.1146/annurev-publhealth-040617-014208](https://doi.org/10.1146/annurev-publhealth-040617-014208)] [Medline: [29261408](https://pubmed.ncbi.nlm.nih.gov/29261408/)]
35. Wesson P, Hswen Y, Valdes G, Stojanovski K, Handley MA. Risks and opportunities to ensure equity in the application of big data research in public health. *Annu Rev Public Health*. Apr 05, 2022;43(1):59-78. [FREE Full text] [doi: [10.1146/annurev-publhealth-051920-110928](https://doi.org/10.1146/annurev-publhealth-051920-110928)] [Medline: [34871504](https://pubmed.ncbi.nlm.nih.gov/34871504/)]
36. Iyamu I, Gómez-Ramírez O, Xu AX, Chang H, Watt S, Mckee G, et al. Challenges in the development of digital public health interventions and mapped solutions: findings from a scoping review. *Digit Health*. 2022;8:20552076221102255. [doi: [10.1177/20552076221102255](https://doi.org/10.1177/20552076221102255)] [Medline: [35656283](https://pubmed.ncbi.nlm.nih.gov/35656283/)]
37. Iyamu I, McKee G, Haag D, Gilbert M. Defining the role of digital public health in the evolving digital health landscape: policy and practice implications in Canada. *Health Promot Chronic Dis Prev Can*. Feb 2024;44(2):66-69. [FREE Full text] [doi: [10.24095/hpcdp.44.2.04](https://doi.org/10.24095/hpcdp.44.2.04)] [Medline: [38353941](https://pubmed.ncbi.nlm.nih.gov/38353941/)]
38. Khan S, Kandukuri DK, Subramaniyan EU, MohanaSundaram A. Harnessing the untapped potential of digital twin technology in digital public health interventions. *Explor Digit Health Technol*. Nov 1, 2023;5(4):11-16. [doi: [10.37349/edht.2023.00003](https://doi.org/10.37349/edht.2023.00003)]
39. Maaß L, Pan CC, Freye M. Mapping digital public health interventions among existing digital technologies and internet-based interventions to maintain and improve population health in practice: protocol for a scoping review. *JMIR Res Protoc*. Mar 31, 2022;11(3):e33404. [FREE Full text] [doi: [10.2196/33404](https://doi.org/10.2196/33404)] [Medline: [35357321](https://pubmed.ncbi.nlm.nih.gov/35357321/)]
40. Odone A, Buttigieg S, Ricciardi W, Azzopardi-Muscat N, Staines A. Public health digitalization in Europe. *Eur J Public Health*. Oct 01, 2019;29(Supplement_3):28-35. [FREE Full text] [doi: [10.1093/eurpub/ckz161](https://doi.org/10.1093/eurpub/ckz161)] [Medline: [31738441](https://pubmed.ncbi.nlm.nih.gov/31738441/)]
41. Wienert J, Janel T, Maaß L. What are digital public health interventions? First steps toward a definition and an intervention classification framework. *J Med Internet Res*. Jun 28, 2022;24(6):e31921. [FREE Full text] [doi: [10.2196/31921](https://doi.org/10.2196/31921)] [Medline: [35763320](https://pubmed.ncbi.nlm.nih.gov/35763320/)]

42. Burgun A, Bernal-Delgado E, Kuchinke W, van Staa T, Cunningham J, Lettieri E, et al. Health data for public health: towards new ways of combining data sources to support research efforts in Europe. *Yearb Med Inform.* Aug 19, 2017;26(01):235-240. [doi: [10.1055/S-0037-1606509](https://doi.org/10.1055/S-0037-1606509)]
43. Canfell OJ, Davidson K, Sullivan C, Eakin E, Burton-Jones A. Data sources for precision public health of obesity: a scoping review, evidence map and use case in Queensland, Australia. *BMC Public Health.* Mar 24, 2022;22(1):584. [FREE Full text] [doi: [10.1186/s12889-022-12939-x](https://doi.org/10.1186/s12889-022-12939-x)] [Medline: [35331189](https://pubmed.ncbi.nlm.nih.gov/35331189/)]
44. Canfell OJ, Davidson K, Woods L, Sullivan C, Cocoros NM, Klompas M, et al. Precision public health for non-communicable diseases: an emerging strategic roadmap and multinational use cases. *Front. Public Health.* Apr 8, 2022;10:854525. [FREE Full text] [doi: [10.3389/fpubh.2022.854525](https://doi.org/10.3389/fpubh.2022.854525)] [Medline: [35462850](https://pubmed.ncbi.nlm.nih.gov/35462850/)]
45. Sielemann K, Hafner A, Pucker B. The reuse of public datasets in the life sciences: potential risks and rewards. *PeerJ.* 2020;8:e9954. [FREE Full text] [doi: [10.7717/peerj.9954](https://doi.org/10.7717/peerj.9954)] [Medline: [33024631](https://pubmed.ncbi.nlm.nih.gov/33024631/)]
46. Deal JA, Betz J, Lin FR, Reed NS. Interpreting results from epidemiologic studies. *Semin Hear.* Feb 15, 2021;42(1):3-9. [FREE Full text] [doi: [10.1055/s-0041-1725996](https://doi.org/10.1055/s-0041-1725996)] [Medline: [33883787](https://pubmed.ncbi.nlm.nih.gov/33883787/)]
47. DiPietro NA. Methods in epidemiology: observational study designs. *Pharmacotherapy.* Oct 06, 2010;30(10):973-984. [doi: [10.1592/phco.30.10.973](https://doi.org/10.1592/phco.30.10.973)] [Medline: [20874034](https://pubmed.ncbi.nlm.nih.gov/20874034/)]
48. Cancer epidemiology: principles and methods. International Agency for Research on Cancer. 1999. URL: <https://publications.iarc.fr/publications/media/download/4106/a72983d2524d630f6ad0539f747e4780447f1ae3.pdf> [accessed 2024-04-29]
49. Munnangi S, Boktor SW. Epidemiology of study design. *StatPearls.* 2014. URL: <https://www.ncbi.nlm.nih.gov/books/NBK470342/> [accessed 2024-04-29]
50. Röhrig B, du Prel JB, Wachtlin D, Blettner M. Types of study in medical research: part 3 of a series on evaluation of scientific publications. *Dtsch Arztebl Int.* Apr 2009;106(15):262-268. [FREE Full text] [doi: [10.3238/arztebl.2009.0262](https://doi.org/10.3238/arztebl.2009.0262)] [Medline: [19547627](https://pubmed.ncbi.nlm.nih.gov/19547627/)]
51. Himali JJ, Baril AA, Cavuoto MG, Yiallourou S, Wiedner CD, Himali D, et al. Association between slow-wave sleep loss and incident dementia. *JAMA Neurol.* Dec 01, 2023;80(12):1326-1333. [doi: [10.1001/jamaneurol.2023.3889](https://doi.org/10.1001/jamaneurol.2023.3889)] [Medline: [37902739](https://pubmed.ncbi.nlm.nih.gov/37902739/)]
52. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. *Lancet.* Mar 15, 2014;383(9921):999-1008. [FREE Full text] [doi: [10.1016/S0140-6736\(13\)61752-3](https://doi.org/10.1016/S0140-6736(13)61752-3)] [Medline: [24084292](https://pubmed.ncbi.nlm.nih.gov/24084292/)]
53. Salines G, De LC. Les cohortes : intérêt, rôle et position de l'InVS. In: Saint-Maurice (Fra) : Institut de veille sanitaire, janvier 2010, 20 p. Princeton, NJ. Princeton University Press; 2018.
54. Zins M, Goldberg M. L'apport des bases de données d'origine administrative aux cohortes épidémiologiques : l'exemple de la cohorte Constances. *Statistique et Société.* 2014;2(2):43-48. [FREE Full text]
55. Chen W, Wang X, Chen J, You C, Ma L, Zhang W, et al. Household air pollution, adherence to a healthy lifestyle, and risk of cardiometabolic multimorbidity: results from the China health and retirement longitudinal study. *Sci Total Environ.* Jan 10, 2023;855:158896. [doi: [10.1016/j.scitotenv.2022.158896](https://doi.org/10.1016/j.scitotenv.2022.158896)] [Medline: [36150596](https://pubmed.ncbi.nlm.nih.gov/36150596/)]
56. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). *Int J Epidemiol.* Feb 12, 2014;43(1):61-68. [FREE Full text] [doi: [10.1093/ije/dys203](https://doi.org/10.1093/ije/dys203)] [Medline: [23243115](https://pubmed.ncbi.nlm.nih.gov/23243115/)]
57. Kesse-Guyot E, Assmann K, Andreeva V, Castetbon K, Méjean C, Touvier M, et al. Metadata correction of: lessons learned from methodological validation research in e-epidemiology. *JMIR Public Health Surveill.* Oct 31, 2016;2(2):e160. [FREE Full text] [doi: [10.2196/publichealth.6876](https://doi.org/10.2196/publichealth.6876)] [Medline: [27797322](https://pubmed.ncbi.nlm.nih.gov/27797322/)]
58. Quantin C. Bases de données médicoadministratives et épidémiologie : intérêt, limites. *Revue d'Épidémiologie et de Santé Publique.* Sep 2008;56(5):S254. [doi: [10.1016/j.respe.2008.06.011](https://doi.org/10.1016/j.respe.2008.06.011)]
59. Babik KR, Shockey TM, Moore LL, Wurzelbacher SJ. Standardizing industrial hygiene data collection forms used by workers' compensation insurers. *J Occup Environ Hyg.* Sep 04, 2018;15(9):676-685. [FREE Full text] [doi: [10.1080/15459624.2018.1490022](https://doi.org/10.1080/15459624.2018.1490022)] [Medline: [29985777](https://pubmed.ncbi.nlm.nih.gov/29985777/)]
60. Bosward R, Braunack-Mayer A, Frost E, Carter S. Mapping precision public health definitions, terminology and applications: a scoping review protocol. *BMJ Open.* Feb 23, 2022;12(2):e058069. [FREE Full text] [doi: [10.1136/bmjopen-2021-058069](https://doi.org/10.1136/bmjopen-2021-058069)] [Medline: [35197357](https://pubmed.ncbi.nlm.nih.gov/35197357/)]
61. Canfell OJ, Kodiyattu Z, Eakin E, Burton-Jones A, Wong I, Macaulay C, et al. Real-world data for precision public health of noncommunicable diseases: a scoping review. *BMC Public Health.* Nov 24, 2022;22(1):2166. [FREE Full text] [doi: [10.1186/s12889-022-14452-7](https://doi.org/10.1186/s12889-022-14452-7)] [Medline: [36434553](https://pubmed.ncbi.nlm.nih.gov/36434553/)]
62. Chao K, Sarker MN, Ali I, Firdaus RB, Azman A, Shaed MM. Big data-driven public health policy making: potential for the healthcare industry. *Heliyon.* Sep 2023;9(9):e19681. [FREE Full text] [doi: [10.1016/j.heliyon.2023.e19681](https://doi.org/10.1016/j.heliyon.2023.e19681)] [Medline: [37809720](https://pubmed.ncbi.nlm.nih.gov/37809720/)]
63. Dolley S. Big Data's role in precision public health. *Front Public Health.* Mar 7, 2018;6:68. [FREE Full text] [doi: [10.3389/fpubh.2018.00068](https://doi.org/10.3389/fpubh.2018.00068)] [Medline: [29594091](https://pubmed.ncbi.nlm.nih.gov/29594091/)]
64. Doyle CM, Lix LM, Hemmelgarn BR, Paterson JM, Renoux C. Data variability across Canadian administrative health databases: differences in content, coding, and completeness. *Pharmacoepidemiol Drug Saf.* Jan 10, 2020;29 Suppl 1(S1):68-77. [doi: [10.1002/pds.4889](https://doi.org/10.1002/pds.4889)] [Medline: [31507029](https://pubmed.ncbi.nlm.nih.gov/31507029/)]

65. Gallini A, Moisan F, Maura G, Carcaillon-Bentata L, Leray E, Haesebaert J, et al. [Identification of neurodegenerative diseases in administrative databases in France: a systematic review of the literature]. *Rev Epidemiol Sante Publique*. Oct 2017;65 Suppl 4:S183-S197. [doi: [10.1016/j.respe.2017.01.115](https://doi.org/10.1016/j.respe.2017.01.115)] [Medline: [28341166](https://pubmed.ncbi.nlm.nih.gov/28341166/)]
66. Khoury MJ, Iademarco MF, Riley WT. Precision public health for the era of precision medicine. *Am J Prev Med*. Mar 2016;50(3):398-401. [FREE Full text] [doi: [10.1016/j.amepre.2015.08.031](https://doi.org/10.1016/j.amepre.2015.08.031)] [Medline: [26547538](https://pubmed.ncbi.nlm.nih.gov/26547538/)]
67. McCullough LE, Maliniak ML, Amin AB, Baker JM, Baliashvili D, Barberio J, et al. Epidemiology beyond its limits. *Sci Adv*. Jun 10, 2022;8(23):eabn3328. [FREE Full text] [doi: [10.1126/sciadv.abn3328](https://doi.org/10.1126/sciadv.abn3328)] [Medline: [35675391](https://pubmed.ncbi.nlm.nih.gov/35675391/)]
68. Näher AF, Vorisek CN, Klopfenstein SA, Lehne M, Thun S, Alsalamah S, et al. Secondary data for global health digitalisation. *Lancet Digit Health*. Feb 2023;5(2):e93-101. [FREE Full text] [doi: [10.1016/S2589-7500\(22\)00195-9](https://doi.org/10.1016/S2589-7500(22)00195-9)] [Medline: [36707190](https://pubmed.ncbi.nlm.nih.gov/36707190/)]
69. Koesten L, Vougiouklis P, Simperl E, Groth P. Dataset reuse: toward translating principles to practice. *Patterns (N Y)*. Nov 13, 2020;1(8):100136. [FREE Full text] [doi: [10.1016/j.patter.2020.100136](https://doi.org/10.1016/j.patter.2020.100136)] [Medline: [33294873](https://pubmed.ncbi.nlm.nih.gov/33294873/)]
70. Hu H, Galea S, Rosella L, Henry D. Big data and population health: focusing on the health impacts of the social, physical, and economic environment. *Epidemiology*. Nov 2017;28(6):759-762. [doi: [10.1097/EDE.0000000000000711](https://doi.org/10.1097/EDE.0000000000000711)] [Medline: [28682850](https://pubmed.ncbi.nlm.nih.gov/28682850/)]
71. Shaw RJ, Harron KL, Pescarini JM, Pinto Junior EP, Allik M, Siroky AN, et al. Biases arising from linked administrative data for epidemiological research: a conceptual framework from registration to analyses. *Eur J Epidemiol*. Dec 05, 2022;37(12):1215-1224. [FREE Full text] [doi: [10.1007/s10654-022-00934-w](https://doi.org/10.1007/s10654-022-00934-w)] [Medline: [36333542](https://pubmed.ncbi.nlm.nih.gov/36333542/)]
72. Subramanian M, Wojtusciszyn A, Favre L, Boughorbel S, Shan J, Letaief KB, et al. Precision medicine in the era of artificial intelligence: implications in chronic disease management. *J Transl Med*. Dec 09, 2020;18(1):472. [FREE Full text] [doi: [10.1186/s12967-020-02658-5](https://doi.org/10.1186/s12967-020-02658-5)] [Medline: [33298113](https://pubmed.ncbi.nlm.nih.gov/33298113/)]
73. Velmovitsky PE, Bevilacqua T, Alencar P, Cowan D, Morita PP. Convergence of precision medicine and public health into precision public health: toward a big data perspective. *Front Public Health*. Apr 6, 2021;9:561873. [FREE Full text] [doi: [10.3389/fpubh.2021.561873](https://doi.org/10.3389/fpubh.2021.561873)] [Medline: [33889555](https://pubmed.ncbi.nlm.nih.gov/33889555/)]
74. Zhang H, Hu H, Diller M, Hogan WR, Prospero M, Guo Y, et al. Semantic standards of external exposome data. *Environ Res*. Jun 2021;197:111185. [FREE Full text] [doi: [10.1016/j.envres.2021.111185](https://doi.org/10.1016/j.envres.2021.111185)] [Medline: [33901445](https://pubmed.ncbi.nlm.nih.gov/33901445/)]
75. Daniels K, Bonnechère B. Harnessing digital health interventions to bridge the gap in prevention for older adults. *Front Public Health*. Jan 8, 2023;11:1281923. [FREE Full text] [doi: [10.3389/fpubh.2023.1281923](https://doi.org/10.3389/fpubh.2023.1281923)] [Medline: [38259780](https://pubmed.ncbi.nlm.nih.gov/38259780/)]
76. Liu F, Panagiotakos D. Real-world data: a brief review of the methods, applications, challenges and opportunities. *BMC Med Res Methodol*. Nov 05, 2022;22(1):287. [FREE Full text] [doi: [10.1186/s12874-022-01768-6](https://doi.org/10.1186/s12874-022-01768-6)] [Medline: [36335315](https://pubmed.ncbi.nlm.nih.gov/36335315/)]
77. Maaß L, Freye M, Pan CC, Dassow HH, Niess J, Jahnel T. The definitions of health apps and medical apps from the perspective of public health and law: qualitative analysis of an interdisciplinary literature overview. *JMIR Mhealth Uhealth*. Oct 31, 2022;10(10):e37980. [FREE Full text] [doi: [10.2196/37980](https://doi.org/10.2196/37980)] [Medline: [36315221](https://pubmed.ncbi.nlm.nih.gov/36315221/)]
78. Mooney SJ, Garber MD. Sampling and sampling frames in big data epidemiology. *Curr Epidemiol Rep*. Mar 2, 2019;6(1):14-22. [FREE Full text] [doi: [10.1007/s40471-019-0179-y](https://doi.org/10.1007/s40471-019-0179-y)] [Medline: [31360626](https://pubmed.ncbi.nlm.nih.gov/31360626/)]
79. Pearson AL, Tribby C, Brown CD, Yang JA, Pfeiffer K, Jankowska MM. Systematic review of best practices for GPS data usage, processing, and linkage in health, exposure science and environmental context research. *BMJ Open*. Feb 01, 2024;14(2):e077036. [FREE Full text] [doi: [10.1136/bmjopen-2023-077036](https://doi.org/10.1136/bmjopen-2023-077036)] [Medline: [38307539](https://pubmed.ncbi.nlm.nih.gov/38307539/)]
80. Prospero M, Min JS, Bian J, Modave F. Big data hurdles in precision medicine and precision public health. *BMC Med Inform Decis Mak*. Dec 29, 2018;18(1):139. [FREE Full text] [doi: [10.1186/s12911-018-0719-2](https://doi.org/10.1186/s12911-018-0719-2)] [Medline: [30594159](https://pubmed.ncbi.nlm.nih.gov/30594159/)]
81. Safran C. Update on data reuse in health care. *Yearb Med Inform*. Aug 2017;26(1):24-27. [FREE Full text] [doi: [10.15265/IY-2017-013](https://doi.org/10.15265/IY-2017-013)] [Medline: [29063535](https://pubmed.ncbi.nlm.nih.gov/29063535/)]
82. Severo M, Feredj A, Romele A. Soft data and public policy: can social media offer alternatives to official statistics in urban policymaking? *Policy Internet*. Jul 19, 2016;8(3):354-372. [doi: [10.1002/poi3.127](https://doi.org/10.1002/poi3.127)]
83. Shakeri Hossein Abad Z, Butler GP, Thompson W, Lee J. Physical activity, sedentary behavior, and sleep on Twitter: multicountry and fully labeled public data set for digital public health surveillance research. *JMIR Public Health Surveill*. Feb 14, 2022;8(2):e32355. [FREE Full text] [doi: [10.2196/32355](https://doi.org/10.2196/32355)] [Medline: [35156938](https://pubmed.ncbi.nlm.nih.gov/35156938/)]
84. Singh T, Roberts K, Cohen T, Cobb N, Wang J, Fujimoto K, et al. Social media as a research tool (SMaaRT) for risky behavior analytics: methodological review. *JMIR Public Health Surveill*. Nov 30, 2020;6(4):e21660. [FREE Full text] [doi: [10.2196/21660](https://doi.org/10.2196/21660)] [Medline: [33252345](https://pubmed.ncbi.nlm.nih.gov/33252345/)]
85. Zins M, Cuggia M, Goldberg M. Les données de santé en France - Abondantes mais complexes. *Med Sci (Paris)*. Feb 16, 2021;37(2):179-184. [FREE Full text] [doi: [10.1051/medsci/2021001](https://doi.org/10.1051/medsci/2021001)] [Medline: [33591261](https://pubmed.ncbi.nlm.nih.gov/33591261/)]
86. Motahari-Nezhad H, Fgaier M, Mahdi Abid M, Péntek M, Gulácsi L, Zrubka Z. Digital biomarker-based studies: scoping review of systematic reviews. *JMIR Mhealth Uhealth*. Oct 24, 2022;10(10):e35722. [FREE Full text] [doi: [10.2196/35722](https://doi.org/10.2196/35722)] [Medline: [36279171](https://pubmed.ncbi.nlm.nih.gov/36279171/)]
87. Ayilara OF, Platt RW, Dahl M, Coulombe J, Ginestet PG, Chateau D, et al. Generating synthetic data from administrative health records for drug safety and effectiveness studies. *Int J Popul Data Sci*. Nov 27, 2023;8(1):2176. [FREE Full text] [doi: [10.23889/ijpds.v8i1.2176](https://doi.org/10.23889/ijpds.v8i1.2176)] [Medline: [38414538](https://pubmed.ncbi.nlm.nih.gov/38414538/)]

88. Fadel M, Petot J, Gourraud PA, Descatha A. Flexibility of a large blindly synthesized avatar database for occupational research: example from the CONSTANCES cohort for stroke and knee pain. *PLoS One*. Jul 31, 2024;19(7):e0308063. [FREE Full text] [doi: [10.1371/journal.pone.0308063](https://doi.org/10.1371/journal.pone.0308063)] [Medline: [39083487](https://pubmed.ncbi.nlm.nih.gov/39083487/)]
89. Giuffrè M, Shung DL. Harnessing the power of synthetic data in healthcare: innovation, application, and privacy. *NPJ Digit Med*. Oct 09, 2023;6(1):186. [FREE Full text] [doi: [10.1038/s41746-023-00927-3](https://doi.org/10.1038/s41746-023-00927-3)] [Medline: [37813960](https://pubmed.ncbi.nlm.nih.gov/37813960/)]
90. Kamel Boulos MN, Zhang P. Digital twins: from personalised medicine to precision public health. *J Pers Med*. Jul 29, 2021;11(8):745. [FREE Full text] [doi: [10.3390/jpm11080745](https://doi.org/10.3390/jpm11080745)] [Medline: [34442389](https://pubmed.ncbi.nlm.nih.gov/34442389/)]
91. Vallevik VB, Babic A, Marshall SE, Elvatun S, Brøgger HB, Alagaratnam S, et al. Can I trust my fake data - a comprehensive quality assessment framework for synthetic tabular data in healthcare. *Int J Med Inform*. May 2024;185:105413. [FREE Full text] [doi: [10.1016/j.ijmedinf.2024.105413](https://doi.org/10.1016/j.ijmedinf.2024.105413)] [Medline: [38493547](https://pubmed.ncbi.nlm.nih.gov/38493547/)]
92. Chamot S, Manaouil C, Fantoni S, Frimat P, Al-Salameh A, Desailly R, et al. Système national des données de santé (SNDS) : perspectives en santé au travail et environnementale. *Arch Mal Prof Environ*. Nov 2023;84(6):101875. [doi: [10.1016/j.admp.2023.101875](https://doi.org/10.1016/j.admp.2023.101875)]
93. Cohen S, Gilutz H, Marelli AJ, Iserin L, Benis A, Bonnet D, et al. Administrative health databases for addressing emerging issues in adults with CHD: a systematic review. *Cardiol Young*. Jun 29, 2018;28(6):844-853. [doi: [10.1017/S1047951118000446](https://doi.org/10.1017/S1047951118000446)] [Medline: [29704902](https://pubmed.ncbi.nlm.nih.gov/29704902/)]
94. Fraser HS, Mugisha M, Bacher I, Ngenzi JL, Seebregts C, Umubyeyi A, et al. Factors influencing data quality in electronic health record systems in 50 health facilities in Rwanda and the role of clinical alerts: cross-sectional observational study. *JMIR Public Health Surveill*. Jul 03, 2024;10:e49127. [FREE Full text] [doi: [10.2196/49127](https://doi.org/10.2196/49127)] [Medline: [38959048](https://pubmed.ncbi.nlm.nih.gov/38959048/)]
95. Gavrielov-Yusim N, Friger M. Use of administrative medical databases in population-based research. *J Epidemiol Community Health*. Mar 2014;68(3):283-287. [doi: [10.1136/jech-2013-202744](https://doi.org/10.1136/jech-2013-202744)] [Medline: [24248997](https://pubmed.ncbi.nlm.nih.gov/24248997/)]
96. Hinds A, Lix LM, Smith M, Quan H, Sanmartin C. Quality of administrative health databases in Canada: a scoping review. *Can J Public Health*. Jun 27, 2016;107(1):e56-e61. [FREE Full text] [doi: [10.17269/cjph.107.5244](https://doi.org/10.17269/cjph.107.5244)] [Medline: [27348111](https://pubmed.ncbi.nlm.nih.gov/27348111/)]
97. Howell CR, Zhang L, Clay OJ, Dutton G, Horton T, Mugavero MJ, et al. Social determinants of health phenotypes and cardiometabolic condition prevalence among patients in a large academic health system: latent class analysis. *JMIR Public Health Surveill*. Aug 07, 2024;10:e53371. [FREE Full text] [doi: [10.2196/53371](https://doi.org/10.2196/53371)] [Medline: [39113389](https://pubmed.ncbi.nlm.nih.gov/39113389/)]
98. Jones KH, Heys S, Tingay KS, Jackson P, Dibben C. The good, the bad, the clunky: improving the use of administrative data for research. *Int J Popul Data Sci*. Dec 04, 2018;4(1):587. [FREE Full text] [doi: [10.23889/ijpds.v4i1.587](https://doi.org/10.23889/ijpds.v4i1.587)] [Medline: [32935024](https://pubmed.ncbi.nlm.nih.gov/32935024/)]
99. Namulanda G, Qualters J, Vaidyanathan A, Roberts E, Richardson M, Fraser A, et al. Electronic health record case studies to advance environmental public health tracking. *J Biomed Inform*. Mar 2018;79:98-104. [FREE Full text] [doi: [10.1016/j.jbi.2018.02.012](https://doi.org/10.1016/j.jbi.2018.02.012)] [Medline: [29476967](https://pubmed.ncbi.nlm.nih.gov/29476967/)]
100. Nicholls SG, Quach P, von Elm E, Guttman A, Moher D, Petersen I, et al. The REporting of studies conducted using observational routinely-collected health data (RECORD) statement: methods for arriving at consensus and developing reporting guidelines. *PLoS One*. 2015;10(5):e0125620. [FREE Full text] [doi: [10.1371/journal.pone.0125620](https://doi.org/10.1371/journal.pone.0125620)] [Medline: [25965407](https://pubmed.ncbi.nlm.nih.gov/25965407/)]
101. Silverwood R, Rajah N, Calderwood L, De Stavola B, Harron K, Ploubidis G. Examining the quality and population representativeness of linked survey and administrative data: guidance and illustration using linked 1958 National Child Development Study and Hospital Episode Statistics data. *Int J Popul Data Sci*. Jan 09, 2024;9(1):2137. [FREE Full text] [doi: [10.23889/ijpds.v9i1.2137](https://doi.org/10.23889/ijpds.v9i1.2137)] [Medline: [38425790](https://pubmed.ncbi.nlm.nih.gov/38425790/)]
102. Zuiderwijk A, Türk BO, Brazier F. Identifying the most important facilitators of open research data sharing and reuse in Epidemiology: a mixed-methods study. *PLoS One*. Feb 8, 2024;19(2):e0297969. [FREE Full text] [doi: [10.1371/journal.pone.0297969](https://doi.org/10.1371/journal.pone.0297969)] [Medline: [38330007](https://pubmed.ncbi.nlm.nih.gov/38330007/)]
103. Hobensack M, Song J, Scharp D, Bowles KH, Topaz M. Machine learning applied to electronic health record data in home healthcare: a scoping review. *Int J Med Inform*. Feb 2023;170:104978. [FREE Full text] [doi: [10.1016/j.ijmedinf.2022.104978](https://doi.org/10.1016/j.ijmedinf.2022.104978)] [Medline: [36592572](https://pubmed.ncbi.nlm.nih.gov/36592572/)]
104. Tassi MF, le Meur N, Stéfic K, Grammatico-Guillon L. Performance of French medico-administrative databases in epidemiology of infectious diseases: a scoping review. *Front Public Health*. May 12, 2023;11:1161550. [FREE Full text] [doi: [10.3389/fpubh.2023.1161550](https://doi.org/10.3389/fpubh.2023.1161550)] [Medline: [37250067](https://pubmed.ncbi.nlm.nih.gov/37250067/)]
105. Yoshida S, Takeuchi M, Tanaka-Mizuno S, Mizuno K, Nakashima M, Fukasawa T, et al. Clinical epidemiology and pharmacoepidemiology studies with real-world databases. *Proc Jpn Acad Ser B Phys Biol Sci*. 2022;98(10):517-528. [FREE Full text] [doi: [10.2183/pjab.98.026](https://doi.org/10.2183/pjab.98.026)] [Medline: [36504194](https://pubmed.ncbi.nlm.nih.gov/36504194/)]
106. Howren A, Aviña-Zubieta JA, Puyat JH, Esdaile JM, Da Costa D, De Vera MA. Defining depression and anxiety in individuals with rheumatic diseases using administrative health databases: a systematic review. *Arthritis Care Res (Hoboken)*. Feb 09, 2020;72(2):243-255. [doi: [10.1002/acr.24048](https://doi.org/10.1002/acr.24048)] [Medline: [31421021](https://pubmed.ncbi.nlm.nih.gov/31421021/)]
107. Kim MK, Roupheal C, McMichael J, Welch N, Dasarathy S. Challenges in and opportunities for electronic health record-based data analysis and interpretation. *Gut Liver*. Mar 15, 2024;18(2):201-208. [FREE Full text] [doi: [10.5009/gnl230272](https://doi.org/10.5009/gnl230272)] [Medline: [37905424](https://pubmed.ncbi.nlm.nih.gov/37905424/)]

108. Lee S, Xu Y, D Apos Souza AG, Martin EA, Doktorchik C, Zhang Z, et al. Unlocking the potential of electronic health records for health research. *Int J Popul Data Sci.* Jan 30, 2020;5(1):1123. [FREE Full text] [doi: [10.23889/ijpds.v5i1.1123](https://doi.org/10.23889/ijpds.v5i1.1123)] [Medline: [32935049](https://pubmed.ncbi.nlm.nih.gov/32935049/)]
109. Pletscher F, Mändli Lerch K, Glinz D. Willingness to share anonymised routinely collected clinical health data in Switzerland: a cross-sectional survey. *Swiss Med Wkly.* Jun 20, 2022;152(2324):w30182. [FREE Full text] [doi: [10.4414/sm.w.2022.w30182](https://doi.org/10.4414/sm.w.2022.w30182)] [Medline: [35752970](https://pubmed.ncbi.nlm.nih.gov/35752970/)]
110. Yuan H, Ali MS, Brouwer ES, Girman CJ, Guo JJ, Lund JL, et al. ISPE Comparative Effectiveness Research Special Interest Group. Real-world evidence: what it is and what it can tell us according to the International Society for Pharmacoepidemiology (ISPE) Comparative Effectiveness Research (CER) Special Interest Group (SIG). *Clin Pharmacol Ther.* Aug 07, 2018;104(2):239-241. [doi: [10.1002/cpt.1086](https://doi.org/10.1002/cpt.1086)] [Medline: [29733448](https://pubmed.ncbi.nlm.nih.gov/29733448/)]
111. Andrew A, Zhou J, Gui J, Harrison A, Shi X, Li M, et al. Pesticides applied to crops and amyotrophic lateral sclerosis risk in the U.S. *Neurotoxicology.* Dec 2021;87:128-135. [FREE Full text] [doi: [10.1016/j.neuro.2021.09.004](https://doi.org/10.1016/j.neuro.2021.09.004)] [Medline: [34562505](https://pubmed.ncbi.nlm.nih.gov/34562505/)]
112. Chazard E, Ficheur G, Caron A, Lamer A, Labreuche J, Cuggia M, et al. Secondary use of healthcare structured data: the challenge of domain-knowledge based extraction of features. *Stud Health Technol Inform.* 2018;255:15-19. [Medline: [30306898](https://pubmed.ncbi.nlm.nih.gov/30306898/)]
113. Grath-Lone LM, Jay MA, Blackburn R, Gordon E, Zylbersztejn A, Wiljaars L, et al. What makes administrative data "research-ready"? A systematic review and thematic analysis of published literature. *Int J Popul Data Sci.* 2022;7(1):1718. [FREE Full text] [doi: [10.23889/ijpds.v6i1.1718](https://doi.org/10.23889/ijpds.v6i1.1718)] [Medline: [35520099](https://pubmed.ncbi.nlm.nih.gov/35520099/)]
114. Hashimoto RE, Brodt ED, Skelly AC, Dettori JR. Administrative database studies: goldmine or goose chase? *Evid Based Spine Care J.* Oct 24, 2014;5(2):74-76. [FREE Full text] [doi: [10.1055/s-0034-1390027](https://doi.org/10.1055/s-0034-1390027)] [Medline: [25278880](https://pubmed.ncbi.nlm.nih.gov/25278880/)]
115. Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, et al. The national healthcare system claims databases in France, SNIIRAM and EGB: powerful tools for pharmacoepidemiology. *Pharmacoepidemiol Drug Saf.* Aug 24, 2017;26(8):954-962. [doi: [10.1002/pds.4233](https://doi.org/10.1002/pds.4233)] [Medline: [28544284](https://pubmed.ncbi.nlm.nih.gov/28544284/)]
116. Al-Sahab B, Leviton A, Loddenkemper T, Paneth N, Zhang B. Biases in electronic health records data for generating real-world evidence: an overview. *J Healthc Inform Res.* Mar 14, 2024;8(1):121-139. [doi: [10.1007/s41666-023-00153-2](https://doi.org/10.1007/s41666-023-00153-2)] [Medline: [38273982](https://pubmed.ncbi.nlm.nih.gov/38273982/)]
117. Bastarache L, Brown JS, Cimino JJ, Dorr DA, Embi PJ, Payne PR, et al. Developing real-world evidence from real-world data: transforming raw data into analytical datasets. *Learn Health Syst.* Jan 14, 2022;6(1):e10293. [FREE Full text] [doi: [10.1002/lrh2.10293](https://doi.org/10.1002/lrh2.10293)] [Medline: [35036557](https://pubmed.ncbi.nlm.nih.gov/35036557/)]
118. Fang J, Kapral MK, Richards J, Robertson A, Stamplecoski M, Silver FL. The Registry of Canadian Stroke Network : an evolving methodology. *Acta Neurol Taiwan.* Jun 2011;20(2):77-84. [Medline: [21739386](https://pubmed.ncbi.nlm.nih.gov/21739386/)]
119. Fillinger S, de la Garza L, Peltzer A, Kohlbacher O, Nahnsen S. Challenges of big data integration in the life sciences. *Anal Bioanal Chem.* Oct 28, 2019;411(26):6791-6800. [doi: [10.1007/s00216-019-02074-9](https://doi.org/10.1007/s00216-019-02074-9)] [Medline: [31463515](https://pubmed.ncbi.nlm.nih.gov/31463515/)]
120. Fonteneau L, Le Meur N, Cohen-Akenine A, Pessel C, Brouard C, Delon F, et al. [The use of administrative health databases in infectious disease epidemiology and public health]. *Rev Epidemiol Sante Publique.* Oct 2017;65 Suppl 4:S174-S182. [FREE Full text] [doi: [10.1016/j.respe.2017.03.131](https://doi.org/10.1016/j.respe.2017.03.131)] [Medline: [28624133](https://pubmed.ncbi.nlm.nih.gov/28624133/)]
121. Silenou BC, Avalos M, Helmer C, Berr C, Pariente A, Jacqmin-Gadda H. Health administrative data enrichment using cohort information: comparative evaluation of methods by simulation and application to real data. *PLoS One.* Jan 31, 2019;14(1):e0211118. [FREE Full text] [doi: [10.1371/journal.pone.0211118](https://doi.org/10.1371/journal.pone.0211118)] [Medline: [30703112](https://pubmed.ncbi.nlm.nih.gov/30703112/)]
122. Wang SV, Schneeweiss S. Data checks before registering study protocols for health care database analyses. *JAMA.* May 07, 2024;331(17):1445-1446. [doi: [10.1001/jama.2024.2988](https://doi.org/10.1001/jama.2024.2988)] [Medline: [38587830](https://pubmed.ncbi.nlm.nih.gov/38587830/)]
123. Yamana H, Yamasaki K, Jo T, Yatsushashi H, Yasunaga H. A descriptive analysis of acute viral hepatitis using a database with electronic medical records and claims data. *Ann Clin Epidemiol.* 2023;5(4):107-112. [FREE Full text] [doi: [10.37737/ace.23014](https://doi.org/10.37737/ace.23014)] [Medline: [38504953](https://pubmed.ncbi.nlm.nih.gov/38504953/)]
124. Gianfrancesco MA, Goldstein ND. A narrative review on the validity of electronic health record-based research in epidemiology. *BMC Med Res Methodol.* Oct 27, 2021;21(1):234. [FREE Full text] [doi: [10.1186/s12874-021-01416-5](https://doi.org/10.1186/s12874-021-01416-5)] [Medline: [34706667](https://pubmed.ncbi.nlm.nih.gov/34706667/)]
125. Denisov NS, Kamenskikh EM, Fedorova OS. Trends in population-based studies: molecular and digital epidemiology (review). *Sovrem Tekhnologii Med.* Aug 29, 2022;14(4):60-70. [FREE Full text] [doi: [10.17691/stm2022.14.4.07](https://doi.org/10.17691/stm2022.14.4.07)] [Medline: [37179982](https://pubmed.ncbi.nlm.nih.gov/37179982/)]
126. Hamm NC, Hamad AF, Wall-Wieler E, Roos LL, Plana-Ripoll O, Lix LM. Multigenerational health research using population-based linked databases: an international review. *Int J Popul Data Sci.* Oct 07, 2021;6(1):1686. [FREE Full text] [doi: [10.23889/ijpds.v6i1.1686](https://doi.org/10.23889/ijpds.v6i1.1686)] [Medline: [34734126](https://pubmed.ncbi.nlm.nih.gov/34734126/)]
127. Poulain R, Tarek MF, Beheshti R. Improving fairness in AI models on electronic health records: the case for federated learning methods. *FAccT 23 (2023).* Jun 2023;2023:1599-1608. [FREE Full text] [doi: [10.1145/3593013.3594102](https://doi.org/10.1145/3593013.3594102)] [Medline: [37990734](https://pubmed.ncbi.nlm.nih.gov/37990734/)]
128. Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L, et al. Value of a national administrative database to guide public decisions: from the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM)

- to the système national des données de santé (SNDS) in France. *Rev Epidemiol Sante Publique*. Oct 2017;65 Suppl 4:S149-S167. [doi: [10.1016/j.respe.2017.05.004](https://doi.org/10.1016/j.respe.2017.05.004)] [Medline: [28756037](https://pubmed.ncbi.nlm.nih.gov/28756037/)]
129. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. *Annu Rev Public Health*. 2016;37:61-81. [FREE Full text] [doi: [10.1146/annurev-publhealth-032315-021353](https://doi.org/10.1146/annurev-publhealth-032315-021353)] [Medline: [26667605](https://pubmed.ncbi.nlm.nih.gov/26667605/)]
 130. Chan JM, Carroll MW, Smyth M, Hamilton Z, Evans D, McGrail K, et al. Comparing health administrative and clinical registry data: trends in incidence and prevalence of pediatric inflammatory bowel disease in British Columbia. *Clin Epidemiol*. 2021;13:81-90. [FREE Full text] [doi: [10.2147/CLEPS.292546](https://doi.org/10.2147/CLEPS.292546)] [Medline: [33603489](https://pubmed.ncbi.nlm.nih.gov/33603489/)]
 131. Chowdhury T, Hemmelgarn B. Evidence-based decision-making 6: utilization of administrative databases for health services research. In: Parfrey PS, Barrett BJ, editors. *Clinical Epidemiology: Practice and Methods*. Cham, Switzerland: Springer; 2015:469-484.
 132. Chishtie J, Sapiro N, Wiebe N, Rabatach L, Lorenzetti D, Leung AA, et al. Use of epic electronic health record system for health care research: scoping review. *J Med Internet Res*. Dec 15, 2023;25:e51003. [FREE Full text] [doi: [10.2196/51003](https://doi.org/10.2196/51003)] [Medline: [38100185](https://pubmed.ncbi.nlm.nih.gov/38100185/)]
 133. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. *Annu Rev Public Health*. 2011;32:91-108. [doi: [10.1146/annurev-publhealth-031210-100700](https://doi.org/10.1146/annurev-publhealth-031210-100700)] [Medline: [21219160](https://pubmed.ncbi.nlm.nih.gov/21219160/)]
 134. Mazzali C, Duca P. Use of administrative data in healthcare research. *Intern Emerg Med*. Jun 25, 2015;10(4):517-524. [doi: [10.1007/s11739-015-1213-9](https://doi.org/10.1007/s11739-015-1213-9)] [Medline: [25711312](https://pubmed.ncbi.nlm.nih.gov/25711312/)]
 135. McPheeters ML, Sathe NA, Jerome RN, Carnahan RM. Methods for systematic reviews of administrative database studies capturing health outcomes of interest. *Vaccine*. Dec 30, 2013;31 Suppl 10:K2-K6. [doi: [10.1016/j.vaccine.2013.06.048](https://doi.org/10.1016/j.vaccine.2013.06.048)] [Medline: [24331070](https://pubmed.ncbi.nlm.nih.gov/24331070/)]
 136. Sato S, Yasunaga H. A review of studies using Japanese nationwide administrative claims databases. *Ann Clin Epidemiol*. 2023;5(2):58-64. [FREE Full text] [doi: [10.37737/ace.23008](https://doi.org/10.37737/ace.23008)] [Medline: [38505730](https://pubmed.ncbi.nlm.nih.gov/38505730/)]
 137. Mues KE, Liede A, Liu J, Wetmore J, Zaha R, Bradbury BD, et al. Use of the Medicare database in epidemiologic and health services research: a valuable source of real-world evidence on the older and disabled populations in the US. *Clin Epidemiol*. 2017;9:267-277. [FREE Full text] [doi: [10.2147/CLEPS.105613](https://doi.org/10.2147/CLEPS.105613)] [Medline: [28533698](https://pubmed.ncbi.nlm.nih.gov/28533698/)]
 138. Bernstein CN. Large registry epidemiology in IBD. *Inflamm Bowel Dis*. Nov 2017;23(11):1941-1949. [doi: [10.1097/MIB.0000000000001279](https://doi.org/10.1097/MIB.0000000000001279)] [Medline: [28991858](https://pubmed.ncbi.nlm.nih.gov/28991858/)]
 139. Bloom PP, Tapper EB. The use of administrative data to investigate the population burden of hepatic encephalopathy. *J Clin Med*. Nov 10, 2020;9(11):3620. [FREE Full text] [doi: [10.3390/jcm9113620](https://doi.org/10.3390/jcm9113620)] [Medline: [33182743](https://pubmed.ncbi.nlm.nih.gov/33182743/)]
 140. Hudson M, Tascilar K, Suisa S. Comparative effectiveness research with administrative health data in rheumatoid arthritis. *Nat Rev Rheumatol*. Jun 15, 2016;12(6):358-366. [doi: [10.1038/nrrheum.2016.34](https://doi.org/10.1038/nrrheum.2016.34)] [Medline: [27080692](https://pubmed.ncbi.nlm.nih.gov/27080692/)]
 141. Mitchell E, Waring T, Ahern E, O'Donovan D, O'Reilly D, Bradley DT. Predictors and consequences of homelessness in whole-population observational studies that used administrative data: a systematic review. *BMC Public Health*. Aug 24, 2023;23(1):1610. [FREE Full text] [doi: [10.1186/s12889-023-16503-z](https://doi.org/10.1186/s12889-023-16503-z)] [Medline: [37612701](https://pubmed.ncbi.nlm.nih.gov/37612701/)]
 142. Ulrich EH, So G, Zappitelli M, Chanchlani R. A review on the application and limitations of administrative health care data for the study of acute kidney injury epidemiology and outcomes in children. *Front Pediatr*. Oct 27, 2021;9:742888. [FREE Full text] [doi: [10.3389/fped.2021.742888](https://doi.org/10.3389/fped.2021.742888)] [Medline: [34778133](https://pubmed.ncbi.nlm.nih.gov/34778133/)]
 143. Curl CL, Spivak M, Phinney R, Montrose L. Synthetic pesticides and health in vulnerable populations: agricultural workers. *Curr Environ Health Rep*. Mar 20, 2020;7(1):13-29. [FREE Full text] [doi: [10.1007/s40572-020-00266-5](https://doi.org/10.1007/s40572-020-00266-5)] [Medline: [31960353](https://pubmed.ncbi.nlm.nih.gov/31960353/)]
 144. Demos K, Sazakli E, Jelastopulu E, Charokopos N, Ellul J, Leotsinidis M. Does farming have an effect on health status? A comparison study in west Greece. *Int J Environ Res Public Health*. Feb 26, 2013;10(3):776-792. [FREE Full text] [doi: [10.3390/ijerph10030776](https://doi.org/10.3390/ijerph10030776)] [Medline: [23442558](https://pubmed.ncbi.nlm.nih.gov/23442558/)]
 145. Kachuri L, Harris MA, MacLeod JS, Tjepkema M, Peters PA, Demers PA. Cancer risks in a population-based study of 70,570 agricultural workers: results from the Canadian census health and Environment cohort (CanCHEC). *BMC Cancer*. May 19, 2017;17(1):343. [FREE Full text] [doi: [10.1186/s12885-017-3346-x](https://doi.org/10.1186/s12885-017-3346-x)] [Medline: [28525996](https://pubmed.ncbi.nlm.nih.gov/28525996/)]
 146. Ramos AK, Girdžiūtė L, Starič J, Rautianinen RH. Identifying "vulnerable agricultural populations" at risk for occupational injuries and illnesses: a European perspective. *J Agromedicine*. Jul 27, 2021;26(3):340-345. [doi: [10.1080/1059924X.2020.1771498](https://doi.org/10.1080/1059924X.2020.1771498)] [Medline: [32458756](https://pubmed.ncbi.nlm.nih.gov/32458756/)]
 147. Sturm ET, Castro C, Mendez-Colmenares A, Duffy J, Burzynska AZ, Stallones L, et al. Risk factors for brain health in agricultural work: a systematic review. *Int J Environ Res Public Health*. Mar 13, 2022;19(6):3373. [FREE Full text] [doi: [10.3390/ijerph19063373](https://doi.org/10.3390/ijerph19063373)] [Medline: [35329061](https://pubmed.ncbi.nlm.nih.gov/35329061/)]
 148. Guindo Y, Parent ME, Richard H, Luce D, Barul C. Expert-based assessment of chemical and physical exposures, and organizational factors, in past agricultural jobs. *Environ Res*. Dec 15, 2024;263(Pt 3):120238. [FREE Full text] [doi: [10.1016/j.envres.2024.120238](https://doi.org/10.1016/j.envres.2024.120238)] [Medline: [39461702](https://pubmed.ncbi.nlm.nih.gov/39461702/)]
 149. Komarek AM, De Pinto A, Smith VH. A review of types of risks in agriculture: what we know and what we need to know. *Agric Syst*. Feb 2020;178:102738. [doi: [10.1016/j.agry.2019.102738](https://doi.org/10.1016/j.agry.2019.102738)]

150. Leon ME, Schinasi LH, Lebailly P, Beane Freeman LE, Nordby KC, Ferro G, et al. Pesticide use and risk of non-Hodgkin lymphoid malignancies in agricultural cohorts from France, Norway and the USA: a pooled analysis from the AGRICOH consortium. *Int J Epidemiol*. Oct 01, 2019;48(5):1519-1535. [FREE Full text] [doi: [10.1093/ije/dyz017](https://doi.org/10.1093/ije/dyz017)] [Medline: [30880337](https://pubmed.ncbi.nlm.nih.gov/30880337/)]
151. Pérès K, Matharan F, Allard M, Amieva H, Baldi I, Barberger-Gateau P, et al. Health and aging in elderly farmers: the AMI cohort. *BMC Public Health*. Jul 27, 2012;12(1):558. [FREE Full text] [doi: [10.1186/1471-2458-12-558](https://doi.org/10.1186/1471-2458-12-558)] [Medline: [22839764](https://pubmed.ncbi.nlm.nih.gov/22839764/)]
152. Petit P, Chamot S, Al-Salameh A, Cancé C, Desailoud R, Bonnetterre V. Farming activity and risk of treated thyroid disorders: insights from the TRACTOR project, a nationwide cohort study. *Environ Res*. May 15, 2024;249:118458. [FREE Full text] [doi: [10.1016/j.envres.2024.118458](https://doi.org/10.1016/j.envres.2024.118458)] [Medline: [38365059](https://pubmed.ncbi.nlm.nih.gov/38365059/)]
153. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. Oct 02, 2018;169(7):467-473. [FREE Full text] [doi: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850)] [Medline: [30178033](https://pubmed.ncbi.nlm.nih.gov/30178033/)]
154. The Joanna Briggs Institute Reviewers' Manual 2015 methodology for JBI scoping reviews. The Joanna Briggs Institute. 2015. URL: <https://reben.com.br/revista/wp-content/uploads/2020/10/Scoping.pdf> [accessed 2024-04-03]
155. Beydon M, Pinto S, De Rycke Y, Fautrel B, Mariette X, Seror R, et al. Risk of cancer for patients with rheumatoid arthritis versus general population: a national claims database cohort study. *Lancet Reg Health Eur*. Dec 2023;35:100768. [FREE Full text] [doi: [10.1016/j.lanepe.2023.100768](https://doi.org/10.1016/j.lanepe.2023.100768)] [Medline: [37954220](https://pubmed.ncbi.nlm.nih.gov/37954220/)]
156. Moysan M. Des archives à la base de données : valeur(s) des données en épidémiologie. *Hist Méd Santé*. Dec 15, 2022;(22):151-164. [FREE Full text] [doi: [10.4000/hms.6274](https://doi.org/10.4000/hms.6274)]
157. Syed R, Eden R, Makasi T, Chukwudi I, Mamudu A, Kamalpour M, et al. Digital health data quality issues: systematic review. *J Med Internet Res*. Mar 31, 2023;25:e42615. [FREE Full text] [doi: [10.2196/42615](https://doi.org/10.2196/42615)] [Medline: [37000497](https://pubmed.ncbi.nlm.nih.gov/37000497/)]
158. Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F. An approach for detecting, quantifying, and visualizing the evolution of a research field: a practical application to the Fuzzy Sets Theory field. *J Informetr*. Jan 2011;5(1):146-166. [doi: [10.1016/j.joi.2010.10.002](https://doi.org/10.1016/j.joi.2010.10.002)]
159. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: an overview and guidelines. *J Bus Res*. Sep 2021;133:285-296. [doi: [10.1016/j.jbusres.2021.04.070](https://doi.org/10.1016/j.jbusres.2021.04.070)]
160. Moral-Muñoz JA, Herrera-Viedma E, Santisteban-Espejo A, Cobo MJ. Software tools for conducting bibliometric analysis in science: an up-to-date review. *Prof Inf J*. Jan 19, 2020;29(1). [FREE Full text] [doi: [10.3145/epi.2020.ene.03](https://doi.org/10.3145/epi.2020.ene.03)]
161. Bienert IR, Oliveira RC, Andrade PB, Caramori CA. Bibliometric indexes, databases and impact factors in cardiology. *Rev Bras Cir Cardiovasc*. 2015;30(2):254-259. [FREE Full text] [doi: [10.5935/1678-9741.20150019](https://doi.org/10.5935/1678-9741.20150019)] [Medline: [26107458](https://pubmed.ncbi.nlm.nih.gov/26107458/)]
162. Bihari A, Tripathi S, Deepak A. A review on h-index and its alternative indices. *J Inf Sci*. May 31, 2021;49(3):624-665. [doi: [10.1177/01655515211014478](https://doi.org/10.1177/01655515211014478)]
163. Shao Y, Chien TW, Jang FL. The use of radar plots with the Yk-index to identify which authors contributed the most to the journal of Medicine in 2020 and 2021: a bibliometric analysis. *Medicine (Baltimore)*. Nov 11, 2022;101(45):e31033. [FREE Full text] [doi: [10.1097/MD.00000000000031033](https://doi.org/10.1097/MD.00000000000031033)] [Medline: [36397440](https://pubmed.ncbi.nlm.nih.gov/36397440/)]
164. Kumar S, Kumar S. Trends of collaborative research in journals of oilseeds research (India), 1993-2004. *Indian J Agric Library Inf Services*. 2008;24:90.
165. Wu T, Duan Y, Zhang T, Tian W, Liu H, Deng Y. Research trends in the application of artificial intelligence in oncology: a bibliometric and network visualization study. *Front Biosci (Landmark Ed)*. Aug 31, 2022;27(9):254. [FREE Full text] [doi: [10.31083/j.fbl2709254](https://doi.org/10.31083/j.fbl2709254)] [Medline: [36224012](https://pubmed.ncbi.nlm.nih.gov/36224012/)]
166. Wang B, Zhao P. Worldwide research productivity in the field of back pain: a bibliometric analysis. *Medicine (Baltimore)*. Oct 2018;97(40):e11566. [FREE Full text] [doi: [10.1097/MD.00000000000011566](https://doi.org/10.1097/MD.00000000000011566)] [Medline: [30290588](https://pubmed.ncbi.nlm.nih.gov/30290588/)]
167. Aria M, Cuccurullo C, D'Aniello L, Misuraca M, Spano M. Thematic analysis as a new culturomic tool: the social media coverage on COVID-19 pandemic in Italy. *Sustainability*. Mar 20, 2022;14(6):3643. [doi: [10.3390/su14063643](https://doi.org/10.3390/su14063643)]
168. Petit P, Leroyer A, Chamot S, Fumery M, Bonnetterre V. Farming activities and risk of inflammatory bowel disease: a French nationwide population-based cohort study. *J Crohns Colitis*. Sep 03, 2024;18(9):1415-1429. [doi: [10.1093/ecco-jcc/jjae050](https://doi.org/10.1093/ecco-jcc/jjae050)] [Medline: [38605515](https://pubmed.ncbi.nlm.nih.gov/38605515/)]
169. Montazeri A, Mohammadi S, M Hesari P, Ghaemi M, Riazi H, Sheikhi-Mobarakeh Z. Preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO): a minimum requirements. *Syst Rev*. Dec 15, 2023;12(1):239. [FREE Full text] [doi: [10.1186/s13643-023-02410-2](https://doi.org/10.1186/s13643-023-02410-2)] [Medline: [38102710](https://pubmed.ncbi.nlm.nih.gov/38102710/)]
170. Aria M, Cuccurullo C. bibliometrix : an R-tool for comprehensive science mapping analysis. *J Informetr*. Nov 2017;11(4):959-975. [doi: [10.1016/j.joi.2017.08.007](https://doi.org/10.1016/j.joi.2017.08.007)]
171. Spitzer WO, Hill GB, Chambers LW, Helliwell BE, Murphy HB. The occupation of fishing as a risk factor in cancer of the lip. *N Engl J Med*. Aug 28, 1975;293(9):419-424. [doi: [10.1056/nejm197508282930903](https://doi.org/10.1056/nejm197508282930903)]
172. Alavanja MC, Samanic C, Dosemeci M, Lubin J, Tarone R, Lynch CF, et al. Use of agricultural pesticides and prostate cancer risk in the Agricultural Health Study cohort. *Am J Epidemiol*. May 01, 2003;157(9):800-814. [doi: [10.1093/aje/kwg040](https://doi.org/10.1093/aje/kwg040)] [Medline: [12727674](https://pubmed.ncbi.nlm.nih.gov/12727674/)]
173. Alavanja MC, Dosemeci M, Samanic C, Lubin J, Lynch CF, Knott C, et al. Pesticides and lung cancer risk in the agricultural health study cohort. *Am J Epidemiol*. Nov 01, 2004;160(9):876-885. [doi: [10.1093/aje/kwh290](https://doi.org/10.1093/aje/kwh290)] [Medline: [15496540](https://pubmed.ncbi.nlm.nih.gov/15496540/)]

174. Andersen A, Barlow L, Engeland A, Kjaerheim K, Lyng E, Pukkala E. Work-related cancer in the Nordic countries. *Scand J Work Environ Health*. 1999;25 Suppl 2:1-116. [FREE Full text] [Medline: [10507118](#)]
175. Andreotti G, Koutros S, Hofmann JN, Sandler DP, Lubin JH, Lynch CF, et al. Glyphosate use and cancer incidence in the agricultural health study. *J Natl Cancer Inst*. May 01, 2018;110(5):509-516. [FREE Full text] [doi: [10.1093/jnci/djx233](#)] [Medline: [29136183](#)]
176. Elbaz A, Clavel J, Rathouz PJ, Moisan F, Galanaud J, Delemotte B, et al. Professional exposure to pesticides and Parkinson disease. *Ann Neurol*. Oct 2009;66(4):494-504. [doi: [10.1002/ana.21717](#)] [Medline: [19847896](#)]
177. Engel LS, Hill DA, Hoppin JA, Lubin JH, Lynch CF, Pierce J, et al. Pesticide use and breast cancer risk among farmers' wives in the agricultural health study. *Am J Epidemiol*. Jan 15, 2005;161(2):121-135. [doi: [10.1093/aje/kwi022](#)] [Medline: [15632262](#)]
178. Eriksson M, Karlsson M. Occupational and other environmental factors and multiple myeloma: a population based case-control study. *Br J Ind Med*. Feb 01, 1992;49(2):95-103. [FREE Full text] [doi: [10.1136/oem.49.2.95](#)] [Medline: [1536825](#)]
179. Hemminki K, Saloniemi I, Salonen T, Partanen T, Vainio H. Childhood cancer and parental occupation in Finland. *J Epidemiol Community Health*. Mar 01, 1981;35(1):11-15. [FREE Full text] [doi: [10.1136/jech.35.1.11](#)] [Medline: [7264527](#)]
180. Karjalainen A, Kurppa K, Martikainen R, Klaukka T, Karjalainen J. Work is related to a substantial portion of adult-onset asthma incidence in the Finnish population. *Am J Respir Crit Care Med*. Aug 15, 2001;164(4):565-568. [doi: [10.1164/ajrccm.164.4.2012146](#)] [Medline: [11520716](#)]
181. Karjalainen A, Martikainen R, Klaukka T, Saarinen K, Uitti J. Risk of asthma among Finnish patients with occupational rhinitis. *Chest*. Jan 2003;123(1):283-288. [doi: [10.1378/chest.123.1.283](#)] [Medline: [12527633](#)]
182. Kristensen P, Andersen A, Irgens LM, Laake P, Bye AS. Incidence and risk factors of cancer among men and women in Norwegian agriculture. *Scand J Work Environ Health*. Feb 1996;22(1):14-26. [FREE Full text] [doi: [10.5271/sjweh.104](#)] [Medline: [8685669](#)]
183. Lee WJ, Sandler DP, Blair A, Samanic C, Cross AJ, Alavanja MC. Pesticide use and colorectal cancer risk in the Agricultural Health Study. *Int J Cancer*. Jul 15, 2007;121(2):339-346. [FREE Full text] [doi: [10.1002/ijc.22635](#)] [Medline: [17390374](#)]
184. Leveque C, Elbaz A, Clavel J, Richard F, Vidal JS, Amouyel P, et al. Association between Parkinson's disease and polymorphisms in the nNOS and iNOS genes in a community-based case-control study. *Hum Mol Genet*. Jan 01, 2003;12(1):79-86. [doi: [10.1093/hmg/ddg009](#)] [Medline: [12490535](#)]
185. Mills PK, Yang R. Prostate cancer risk in California farm workers. *J Occup Environ Med*. Mar 2003;45(3):249-258. [doi: [10.1097/01.jom.0000058339.05741.0c](#)] [Medline: [12661182](#)]
186. Morrison H, Savitz D, Semenciw R, Hulka B, Mao Y, Morison D, et al. Farming and prostate cancer mortality. *Am J Epidemiol*. Feb 01, 1993;137(3):270-280. [doi: [10.1093/oxfordjournals.aje.a116674](#)] [Medline: [8452135](#)]
187. Orton SM, Wald L, Confavreux C, Vukusic S, Krohn J, Ramagopalan S, et al. Association of UV radiation with multiple sclerosis prevalence and sex ratio in France. *Neurology*. Feb 2011;76(5):425-431. [doi: [10.1212/wnl.0b013e31820a0a9f](#)]
188. Pouchieu C, Piel C, Carles C, Gruber A, Helmer C, Tual S, et al. Pesticide use in agriculture and Parkinson's disease in the AGRICAN cohort study. *Int J Epidemiol*. Feb 01, 2018;47(1):299-310. [doi: [10.1093/ije/dyx225](#)] [Medline: [29136149](#)]
189. Pukkala E, Martinsen JI, Lyng E, Gunnarsdottir HK, Sparén P, Tryggvadottir L, et al. Occupation and cancer - follow-up of 15 million people in five Nordic countries. *Acta Oncol*. Jul 21, 2009;48(5):646-790. [doi: [10.1080/02841860902913546](#)] [Medline: [19925375](#)]
190. Reif J, Pearce N, Fraser J. Cancer risks in New Zealand farmers. *Int J Epidemiol*. Dec 1989;18(4):768-774. [doi: [10.1093/ije/18.4.768](#)] [Medline: [2621012](#)]
191. Stokes L, Stark A, Marshall E, Narang A. Neurotoxicity among pesticide applicators exposed to organophosphates. *Occup Environ Med*. Oct 01, 1995;52(10):648-653. [FREE Full text] [doi: [10.1136/oem.52.10.648](#)] [Medline: [7489054](#)]
192. Vingård E, Alfredsson L, Goldie I, Hogstedt C. Occupation and osteoarthritis of the hip and knee: a register-based cohort study. *Int J Epidemiol*. Dec 1991;20(4):1025-1031. [doi: [10.1093/ije/20.4.1025](#)] [Medline: [1800399](#)]
193. Vukusic S, Van Bockstael V, Gosselin S, Confavreux C. Regional variations in the prevalence of multiple sclerosis in French farmers. *J Neurol Neurosurg Psychiatry*. Jul 18, 2007;78(7):707-709. [FREE Full text] [doi: [10.1136/jnnp.2006.101196](#)] [Medline: [17299020](#)]
194. Weichenthal S, Villeneuve PJ, Burnett RT, van Donkelaar A, Martin RV, Jones RR, et al. Long-term exposure to fine particulate matter: association with nonaccidental and cardiovascular mortality in the agricultural health study cohort. *Environ Health Perspect*. Jun 2014;122(6):609-615. [FREE Full text] [doi: [10.1289/ehp.1307277](#)] [Medline: [24633320](#)]
195. Wigle DT, Semenciw RM, Wilkins K, Riedel D, Ritter L, Morrison HI, et al. Mortality study of Canadian male farm operators: non-Hodgkin's lymphoma mortality and agricultural practices in Saskatchewan. *J Natl Cancer Inst*. Apr 04, 1990;82(7):575-582. [doi: [10.1093/jnci/82.7.575](#)] [Medline: [2313734](#)]
196. Wiklund K, Dich J, Holm LE, Eklund G. Risk of cancer in pesticide applicators in Swedish agriculture. *Br J Ind Med*. Nov 01, 1989;46(11):809-814. [FREE Full text] [doi: [10.1136/oem.46.11.809](#)] [Medline: [2590647](#)]
197. Horwitz H, Ahlgren B, Nærum E. Effect of occupation on risk of developing MS: an insurance cohort study. *BMJ Open*. Jun 20, 2013;3(6):e002894. [FREE Full text] [doi: [10.1136/bmjopen-2013-002894](#)] [Medline: [23794592](#)]

198. Karttunen JP, Rautiainen RH, Quendler E. Gender division of farm work and occupational injuries. *J Agric Saf Health*. Jul 16, 2019;25(3):117-127. [doi: [10.13031/jash.13177](https://doi.org/10.13031/jash.13177)] [Medline: [32429612](https://pubmed.ncbi.nlm.nih.gov/32429612/)]
199. Keller JE, Howe HL. Risk factors for lung cancer among nonsmoking Illinois residents. *Environ Res*. Jan 1993;60(1):1-11. [doi: [10.1006/enrs.1993.1001](https://doi.org/10.1006/enrs.1993.1001)] [Medline: [8432260](https://pubmed.ncbi.nlm.nih.gov/8432260/)]
200. Marant Micallef C, Shield KD, Vignat J, Baldi I, Charbotel B, Fervers B, et al. Cancers in France in 2015 attributable to occupational exposures. *Int J Hyg Environ Health*. Jan 2019;222(1):22-29. [doi: [10.1016/j.ijheh.2018.07.015](https://doi.org/10.1016/j.ijheh.2018.07.015)] [Medline: [30174219](https://pubmed.ncbi.nlm.nih.gov/30174219/)]
201. Moisan F, Spinosi J, Dupupet JL, Delabre L, Mazurie J, Goldberg M, et al. The relation between type of farming and prevalence of Parkinson's disease among agricultural workers in five French districts. *Mov Disord*. Feb 01, 2011;26(2):271-279. [FREE Full text] [doi: [10.1002/mds.23370](https://doi.org/10.1002/mds.23370)] [Medline: [21412834](https://pubmed.ncbi.nlm.nih.gov/21412834/)]
202. Toch-Marquardt M, Menvielle G, Eikemo TA, Kulhánová I, Kulik MC, Bopp M, et al. Euro-GBD-SE consortium. Occupational class inequalities in all-cause and cause-specific mortality among middle-aged men in 14 European populations during the early 2000s. *PLoS One*. Sep 30, 2014;9(9):e108072. [FREE Full text] [doi: [10.1371/journal.pone.0108072](https://doi.org/10.1371/journal.pone.0108072)] [Medline: [25268702](https://pubmed.ncbi.nlm.nih.gov/25268702/)]
203. Tüchsen F, Andersen O, Costa G, Filakti H, Marmot MG. Occupation and ischemic heart disease in the European Community: a comparative study of occupations at potential high risk. *Am J Ind Med*. Oct 1996;30(4):407-414. [doi: [10.1002/\(SICI\)1097-0274\(199610\)30:4<407::AID-AJIM5>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1097-0274(199610)30:4<407::AID-AJIM5>3.0.CO;2-S)] [Medline: [8892545](https://pubmed.ncbi.nlm.nih.gov/8892545/)]
204. Wang CC, Lin GL, Lin YJ, Chen WL, Wu WT. Occupational health surveillance and detection of emerging occupational diseases among Taiwan farmers, through analysis of national-based farmers' and medico-administrative databases. *Am J Ind Med*. Jan 17, 2023;66(1):85-93. [doi: [10.1002/ajim.23443](https://doi.org/10.1002/ajim.23443)] [Medline: [36398405](https://pubmed.ncbi.nlm.nih.gov/36398405/)]
205. Langholz PL, Sommerseth HL. Old-age mortality and social class in northern Norway in the first half of the twentieth century. *Scand Econ Hist Rev*. Sep 13, 2023;72(2):153-170. [doi: [10.1080/03585522.2023.2255599](https://doi.org/10.1080/03585522.2023.2255599)]
206. Das AV, Bagga B, Joseph J. Clinical profile and demographic distribution of keratitis: an electronic medical record-driven data analytics from an eye care network in India. *Ocul Immunol Inflamm*. Jul 21, 2024;32(5):722-726. [doi: [10.1080/09273948.2023.2199331](https://doi.org/10.1080/09273948.2023.2199331)] [Medline: [37084281](https://pubmed.ncbi.nlm.nih.gov/37084281/)]
207. Lemarchand C, Tual S, Levêque-Morlais N, Perrier S, Belot A, Velten M, et al. AGRICAN group. Cancer incidence in the AGRICAN cohort study (2005-2011). *Cancer Epidemiol*. Aug 2017;49:175-185. [FREE Full text] [doi: [10.1016/j.canep.2017.06.003](https://doi.org/10.1016/j.canep.2017.06.003)] [Medline: [28709062](https://pubmed.ncbi.nlm.nih.gov/28709062/)]
208. Scott E, Hirabayashi L, Levenstein A, Krupa N, Jenkins P. The development of a machine learning algorithm to identify occupational injuries in agriculture using pre-hospital care reports. *Health Inf Sci Syst*. Dec 29, 2021;9(1):31. [FREE Full text] [doi: [10.1007/s13755-021-00161-9](https://doi.org/10.1007/s13755-021-00161-9)] [Medline: [34422257](https://pubmed.ncbi.nlm.nih.gov/34422257/)]
209. Wilke RA, Qamar M, Lupu RA, Gu S, Zhao J. Chronic kidney disease in agricultural communities. *Am J Med*. Oct 2019;132(10):e727-e732. [FREE Full text] [doi: [10.1016/j.amjmed.2019.03.036](https://doi.org/10.1016/j.amjmed.2019.03.036)] [Medline: [30998912](https://pubmed.ncbi.nlm.nih.gov/30998912/)]
210. Engel LS, Werder E, Satagopan J, Blair A, Hoppin JA, Koutros S, et al. Insecticide use and breast cancer risk among farmers' wives in the agricultural health study. *Environ Health Perspect*. Sep 06, 2017;125(9):097002. [FREE Full text] [doi: [10.1289/EHP1295](https://doi.org/10.1289/EHP1295)] [Medline: [28934092](https://pubmed.ncbi.nlm.nih.gov/28934092/)]
211. Louis LM, Lerro CC, Friesen MC, Andreotti G, Koutros S, Sandler DP, et al. A prospective study of cancer risk among agricultural health study farm spouses associated with personal use of organochlorine insecticides. *Environ Health*. Sep 06, 2017;16(1):95. [FREE Full text] [doi: [10.1186/s12940-017-0298-1](https://doi.org/10.1186/s12940-017-0298-1)] [Medline: [28874165](https://pubmed.ncbi.nlm.nih.gov/28874165/)]
212. Elbaek Pedersen J, Hansen J. Risk of breast cancer in daughters of agricultural workers in Denmark. *Environ Res*. Jan 01, 2024;240(Pt 1):117374. [doi: [10.1016/j.envres.2023.117374](https://doi.org/10.1016/j.envres.2023.117374)] [Medline: [37866542](https://pubmed.ncbi.nlm.nih.gov/37866542/)]
213. Kasten-Arias C, Hodes T, Marino M, Kaufmann J, Lucas JA, Estela Vasquez Guzman C, et al. Healthcare utilization for asthma exacerbation among children of migrant and seasonal farmworkers. *Prev Med Rep*. Feb 2024;38:102598. [FREE Full text] [doi: [10.1016/j.pmedr.2024.102598](https://doi.org/10.1016/j.pmedr.2024.102598)] [Medline: [38283959](https://pubmed.ncbi.nlm.nih.gov/38283959/)]
214. Mengist Dessie A, Fenta Feleke S, Getaye Workie S, Getinet Abebe T, Mossu Chanie Y, Kassa Yalew A. Prevalence of skin disease and its associated factors among primary schoolchildren: a cross-sectional study from a Northern Ethiopian town. *Clin Cosmet Investig Dermatol*. 2022;15:791-801. [FREE Full text] [doi: [10.2147/CCID.S361051](https://doi.org/10.2147/CCID.S361051)] [Medline: [35521561](https://pubmed.ncbi.nlm.nih.gov/35521561/)]
215. Mills PK, Zahm SH. Organophosphate pesticide residues in urine of farmworkers and their children in Fresno County, California. *Am J Ind Med*. 2001;40(5):571-577. [doi: [10.1002/ajim.10007.abs](https://doi.org/10.1002/ajim.10007.abs)]
216. Andrew N, Beare R, Ravipati T, Parker E, Snowdon D, Naude K, et al. Developing a linked electronic health record derived data platform to support research into healthy ageing. *Int J Popul Data Sci*. Jun 12, 2023;8(1):2129. [FREE Full text] [doi: [10.23889/ijpds.v8i1.2129](https://doi.org/10.23889/ijpds.v8i1.2129)] [Medline: [37670961](https://pubmed.ncbi.nlm.nih.gov/37670961/)]
217. Contiero P, Borgini A, Bertoldi M, Abita A, Cuffari G, Tomao P, et al. An epidemiological study to investigate links between atmospheric pollution from farming and SARS-CoV-2 mortality. *Int J Environ Res Public Health*. Apr 12, 2022;19(8):4637. [FREE Full text] [doi: [10.3390/ijerph19084637](https://doi.org/10.3390/ijerph19084637)] [Medline: [35457503](https://pubmed.ncbi.nlm.nih.gov/35457503/)]
218. Lotterman A, Baliatsas C, de Rooij MM, Huss A, Jacobs J, Dücker M, et al. Increased risk of pneumonia amongst residents living near goat farms in different livestock-dense regions in the Netherlands. *PLoS One*. Jul 5, 2023;18(7):e0286972. [FREE Full text] [doi: [10.1371/journal.pone.0286972](https://doi.org/10.1371/journal.pone.0286972)] [Medline: [37405987](https://pubmed.ncbi.nlm.nih.gov/37405987/)]

219. Nerich V, Monnet E, Weill A, Vallier N, Vanbockstael V, Auleley G, et al. Fine-scale geographic variations of inflammatory bowel disease in France: correlation with socioeconomic and house equipment variables. *Inflamm Bowel Dis*. May 2010;16(5):813-821. [doi: [10.1002/ibd.21122](https://doi.org/10.1002/ibd.21122)] [Medline: [19902547](https://pubmed.ncbi.nlm.nih.gov/19902547/)]
220. Nilles JD, Lim D, Boyer MP, Wilson BD, Betar RA, Showalter HA, et al. The occurrence of bone and joint cancers and their association with rural living and radon exposure in Iowa. *Environ Geochem Health*. Mar 05, 2023;45(3):925-940. [FREE Full text] [doi: [10.1007/s10653-022-01261-5](https://doi.org/10.1007/s10653-022-01261-5)] [Medline: [35381949](https://pubmed.ncbi.nlm.nih.gov/35381949/)]
221. Yzermans CJ, Moleman YP, Spreeuwenberg P, Nielen MM, Dückers ML, Smit LA, et al. Risk of pneumonia in the vicinity of goat farms: a comparative assessment of temporal variation based on longitudinal health data. *Pneumonia (Nathan)*. Sep 05, 2023;15(1):13. [FREE Full text] [doi: [10.1186/s41479-023-00115-7](https://doi.org/10.1186/s41479-023-00115-7)] [Medline: [37667350](https://pubmed.ncbi.nlm.nih.gov/37667350/)]
222. Petit P, Bosson-Rieutort D, Maugard C, Gondard E, Ozenfant D, Joubert N, et al. The TRACTOR project: TRACKing and MoniToring occupational risks in agriculture using French insurance health data (MSA). *Ann Work Expo Health*. Mar 15, 2022;66(3):402-411. [doi: [10.1093/annweh/wxab083](https://doi.org/10.1093/annweh/wxab083)] [Medline: [34562080](https://pubmed.ncbi.nlm.nih.gov/34562080/)]
223. Mills PK, Kwong S. Cancer incidence in the United Farmworkers of America (UFW), 1987–1997. *Am J Ind Med*. 2001;40(5):596-603. [doi: [10.1002/ajim.1125.abs](https://doi.org/10.1002/ajim.1125.abs)]
224. Guillien A, Laurent L, Soumagne T, Puyraveau M, Laplante J, Andujar P, et al. Anxiety and depression among dairy farmers: the impact of COPD. *Int J Chron Obstruct Pulmon Dis*. 2018;13:1-9. [FREE Full text] [doi: [10.2147/COPD.S143883](https://doi.org/10.2147/COPD.S143883)] [Medline: [29296078](https://pubmed.ncbi.nlm.nih.gov/29296078/)]
225. A consortium of agricultural cohort studies. AGRICOH. URL: <https://agricoh.iarc.fr/> [accessed 2024-05-17]
226. Jouneau S, Pinault M, Gouyet T, Brinchault G, Guillot S, Viel J, et al. GERPRA. [AIRBAg study: preliminary results after one year of screening for COPD in dairy farmers]. *Rev Mal Respir*. Mar 2015;32(3):240-248. [doi: [10.1016/j.rmr.2014.03.014](https://doi.org/10.1016/j.rmr.2014.03.014)] [Medline: [25847201](https://pubmed.ncbi.nlm.nih.gov/25847201/)]
227. Soto S, Yoder AM, Nuño T, Aceves B, Sepulveda R, Rosales CB. Health conditions among farmworkers in the southwest: an analysis of the National Agricultural Workers survey. *Front Public Health*. Nov 3, 2022;10:962085. [FREE Full text] [doi: [10.3389/fpubh.2022.962085](https://doi.org/10.3389/fpubh.2022.962085)] [Medline: [36407981](https://pubmed.ncbi.nlm.nih.gov/36407981/)]
228. Contaxis N, Clark J, Dellureficio A, Gonzales S, Mannheimer S, Oxley PR, et al. Ten simple rules for improving research data discovery. *PLoS Comput Biol*. Feb 10, 2022;18(2):e1009768. [FREE Full text] [doi: [10.1371/journal.pcbi.1009768](https://doi.org/10.1371/journal.pcbi.1009768)] [Medline: [35143479](https://pubmed.ncbi.nlm.nih.gov/35143479/)]
229. Vesteghem C, Brøndum RF, Sønderkær M, Sommer M, Schmitz A, Bødker JS, et al. Implementing the FAIR Data Principles in precision oncology: review of supporting initiatives. *Brief Bioinform*. May 21, 2020;21(3):936-945. [FREE Full text] [doi: [10.1093/bib/bbz044](https://doi.org/10.1093/bib/bbz044)] [Medline: [31263868](https://pubmed.ncbi.nlm.nih.gov/31263868/)]
230. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. Mar 15, 2016;3(1):160018. [FREE Full text] [doi: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)] [Medline: [26978244](https://pubmed.ncbi.nlm.nih.gov/26978244/)]
231. Roux P, Guillien A, Soumagne T, Ritter O, Laplante J, Travers C, et al. Smoking habits in French farmers: a cross-sectional study. *BMC Public Health*. Feb 04, 2017;17(1):166. [FREE Full text] [doi: [10.1186/s12889-017-4030-4](https://doi.org/10.1186/s12889-017-4030-4)] [Medline: [28160776](https://pubmed.ncbi.nlm.nih.gov/28160776/)]
232. Audi C, Baiz N, Maesano CN, Ramousse O, Reboulleau D, Magnan A, et al. Serum cytokine levels related to exposure to volatile organic compounds and PM2.5 in dwellings and workplaces in French farmers - a mechanism to explain nonsmoking COPD. *Int J Chron Obstruct Pulmon Dis*. 2017;12:1363-1374. [FREE Full text] [doi: [10.2147/COPD.S117866](https://doi.org/10.2147/COPD.S117866)] [Medline: [28503065](https://pubmed.ncbi.nlm.nih.gov/28503065/)]
233. Baldi I, Gruber A, Rondeau V, Lebailly P, Brochard P, Fabrigoule C. Neurobehavioral effects of long-term exposure to pesticides: results from the 4-year follow-up of the PHYTONER study. *Occup Environ Med*. Feb 22, 2011;68(2):108-115. [doi: [10.1136/oem.2009.047811](https://doi.org/10.1136/oem.2009.047811)] [Medline: [21097948](https://pubmed.ncbi.nlm.nih.gov/21097948/)]
234. Christensen CH, Platz EA, Andreotti G, Blair A, Hoppin JA, Koutros S, et al. Coumaphos exposure and incident cancer among male participants in the Agricultural Health Study (AHS). *Environ Health Perspect*. Jan 2010;118(1):92-96. [FREE Full text] [doi: [10.1289/ehp.0800446](https://doi.org/10.1289/ehp.0800446)] [Medline: [20056581](https://pubmed.ncbi.nlm.nih.gov/20056581/)]
235. Achard P, Maugard C, Cancé C, Spinosi J, Ozenfant D, Maître A, et al. Medico-administrative data combined with agricultural practices data to retrospectively estimate pesticide use by agricultural workers. *J Expo Sci Environ Epidemiol*. Jul 04, 2020;30(4):743-755. [doi: [10.1038/s41370-019-0166-x](https://doi.org/10.1038/s41370-019-0166-x)] [Medline: [31484997](https://pubmed.ncbi.nlm.nih.gov/31484997/)]
236. Hohenadel K, Harris SA, McLaughlin JR, Spinelli JJ, Pahwa P, Dosman JA, et al. Exposure to multiple pesticides and risk of non-Hodgkin lymphoma in men from six Canadian provinces. *Int J Environ Res Public Health*. Jun 21, 2011;8(6):2320-2330. [FREE Full text] [doi: [10.3390/ijerph8062320](https://doi.org/10.3390/ijerph8062320)] [Medline: [21776232](https://pubmed.ncbi.nlm.nih.gov/21776232/)]
237. Chung EY, Risi D, Holt JL, Lonergan M, Kotwal S, Yong K, et al. Retrospective study on the epidemiology of antineutrophil cytoplasmic autoantibodies-associated vasculitis in two Australian health districts. *Intern Med J*. Apr 13, 2022;52(4):605-613. [doi: [10.1111/imj.15098](https://doi.org/10.1111/imj.15098)] [Medline: [33040456](https://pubmed.ncbi.nlm.nih.gov/33040456/)]
238. Li X, Sundquist J, Sundquist K. Socioeconomic and occupational risk factors for rheumatoid arthritis: a nationwide study based on hospitalizations in Sweden. *J Rheumatol*. Jun 2008;35(6):986-991. [Medline: [18464310](https://pubmed.ncbi.nlm.nih.gov/18464310/)]
239. Apalla Z, Lallas A, Sotiriou E, Lazaridou E, Vakirlis E, Trakatelli M, et al. Farmers develop more aggressive histologic subtypes of basal cell carcinoma. Experience from a tertiary hospital in Northern Greece. *J Eur Acad Dermatol Venereol*. Apr 17, 2016;30 Suppl 3(S3):17-20. [doi: [10.1111/jdv.13605](https://doi.org/10.1111/jdv.13605)] [Medline: [26995018](https://pubmed.ncbi.nlm.nih.gov/26995018/)]

240. Lope V, Pollán M, Gustavsson P, Plato N, Pérez-Gómez B, Aragonés N, et al. Occupation and thyroid cancer risk in Sweden. *J Occup Environ Med*. Sep 2005;47(9):948-957. [doi: [10.1097/01.jom.0000169564.21523.5d](https://doi.org/10.1097/01.jom.0000169564.21523.5d)] [Medline: [16155480](https://pubmed.ncbi.nlm.nih.gov/16155480/)]
241. Nordby KC, Andersen A, Kristensen P. Incidence of lip cancer in the male Norwegian agricultural population. *Cancer Causes Control*. Aug 2004;15(6):619-626. [doi: [10.1023/b:caco.0000036169.90864.e2](https://doi.org/10.1023/b:caco.0000036169.90864.e2)]
242. Kristensen P, Andersen A, Irgens LM. Hormone-dependent cancer and adverse reproductive outcomes in farmers' families--effects of climatic conditions favoring fungal growth in grain. *Scand J Work Environ Health*. Aug 2000;26(4):331-337. [FREE Full text] [doi: [10.5271/sjweh.550](https://doi.org/10.5271/sjweh.550)] [Medline: [10994799](https://pubmed.ncbi.nlm.nih.gov/10994799/)]
243. Døssing M, Petersen KT, Vyberg M, Olsen JH. Liver cancer among employees in Denmark. *Am J Ind Med*. Sep 1997;32(3):248-254. [doi: [10.1002/\(sici\)1097-0274\(199709\)32:3<248::aid-ajim10>3.0.co;2-v](https://doi.org/10.1002/(sici)1097-0274(199709)32:3<248::aid-ajim10>3.0.co;2-v)] [Medline: [9219654](https://pubmed.ncbi.nlm.nih.gov/9219654/)]
244. Klingelschmidt J, Chastang JF, Khireddine-Medouni I, Chérié-Challine L, Niedhammer I. [Occupational factors associated with suicide among French employees from the special agricultural social security scheme (MSA) working between 2007 and 2013]. *Rev Epidemiol Sante Publique*. Feb 2020;68(1):1-8. [doi: [10.1016/j.respe.2019.09.008](https://doi.org/10.1016/j.respe.2019.09.008)] [Medline: [31843361](https://pubmed.ncbi.nlm.nih.gov/31843361/)]
245. Manninen P, Heliövaara M, Riihimäki H, Mäkelä P. Does psychological distress predict disability? *Int J Epidemiol*. Oct 1997;26(5):1063-1070. [doi: [10.1093/ije/26.5.1063](https://doi.org/10.1093/ije/26.5.1063)] [Medline: [9363529](https://pubmed.ncbi.nlm.nih.gov/9363529/)]
246. Zöller B, Li X, Sundquist J, Sundquist K. Socioeconomic and occupational risk factors for venous thromboembolism in Sweden: a nationwide epidemiological study. *Thromb Res*. May 2012;129(5):577-582. [doi: [10.1016/j.thromres.2011.07.050](https://doi.org/10.1016/j.thromres.2011.07.050)] [Medline: [21868069](https://pubmed.ncbi.nlm.nih.gov/21868069/)]
247. Lucyk K, Lu M, Sajobi T, Quan H. Administrative health data in Canada: lessons from history. *BMC Med Inform Decis Mak*. Aug 19, 2015;15(1):69. [FREE Full text] [doi: [10.1186/s12911-015-0196-9](https://doi.org/10.1186/s12911-015-0196-9)] [Medline: [26286712](https://pubmed.ncbi.nlm.nih.gov/26286712/)]
248. Kab S, Spinosi J, Chaperon L, Dugravot A, Singh-Manoux A, Moisan F, et al. Agricultural activities and the incidence of Parkinson's disease in the general French population. *Eur J Epidemiol*. Mar 9, 2017;32(3):203-216. [doi: [10.1007/s10654-017-0229-z](https://doi.org/10.1007/s10654-017-0229-z)] [Medline: [28185034](https://pubmed.ncbi.nlm.nih.gov/28185034/)]
249. Perrin L, Spinosi J, Chaperon L, Kab S, Moisan F, Ebaz A. Pesticides expenditures by farming type and incidence of Parkinson disease in farmers: a French nationwide study. *Environ Res*. Jun 2021;197:111161. [doi: [10.1016/j.envres.2021.111161](https://doi.org/10.1016/j.envres.2021.111161)] [Medline: [33887276](https://pubmed.ncbi.nlm.nih.gov/33887276/)]
250. Carles C, Bouvier G, Lebailly P, Baldi I. Use of job-exposure matrices to estimate occupational exposure to pesticides: a review. *J Expo Sci Environ Epidemiol*. Mar 2017;27(2):125-140. [doi: [10.1038/jes.2016.25](https://doi.org/10.1038/jes.2016.25)] [Medline: [27189257](https://pubmed.ncbi.nlm.nih.gov/27189257/)]
251. Tudi M, Li H, Li H, Wang L, Lyu J, Yang L, et al. Exposure routes and health risks associated with pesticide application. *Toxics*. Jun 19, 2022;10(6):335. [FREE Full text] [doi: [10.3390/toxics10060335](https://doi.org/10.3390/toxics10060335)] [Medline: [35736943](https://pubmed.ncbi.nlm.nih.gov/35736943/)]
252. Ahmad MF, Ahmad FA, Alsayegh AA, Zeyaullah M, AlShahrani AM, Muzammil K, et al. Pesticides impacts on human health and the environment with their mechanisms of action and possible countermeasures. *Heliyon*. Apr 15, 2024;10(7):e29128. [FREE Full text] [doi: [10.1016/j.heliyon.2024.e29128](https://doi.org/10.1016/j.heliyon.2024.e29128)] [Medline: [38623208](https://pubmed.ncbi.nlm.nih.gov/38623208/)]
253. Cavalier H, Trasande L, Porta M. Exposures to pesticides and risk of cancer: evaluation of recent epidemiological evidence in humans and paths forward. *Int J Cancer*. Mar 01, 2023;152(5):879-912. [FREE Full text] [doi: [10.1002/ijc.34300](https://doi.org/10.1002/ijc.34300)] [Medline: [36134639](https://pubmed.ncbi.nlm.nih.gov/36134639/)]
254. Cresto N, Forner-Piquer I, Baig A, Chatterjee M, Perroy J, Goracci J, et al. Pesticides at brain borders: impact on the blood-brain barrier, neuroinflammation, and neurological risk trajectories. *Chemosphere*. May 2023;324:138251. [doi: [10.1016/j.chemosphere.2023.138251](https://doi.org/10.1016/j.chemosphere.2023.138251)] [Medline: [36878369](https://pubmed.ncbi.nlm.nih.gov/36878369/)]
255. Richardson JR, Fitsanakis V, Westerink RH, Kanthasamy AG. Neurotoxicity of pesticides. *Acta Neuropathol*. Sep 13, 2019;138(3):343-362. [FREE Full text] [doi: [10.1007/s00401-019-02033-9](https://doi.org/10.1007/s00401-019-02033-9)] [Medline: [31197504](https://pubmed.ncbi.nlm.nih.gov/31197504/)]
256. Sun J, Fang R, Wang H, Xu D, Yang J, Huang X, et al. A review of environmental metabolism disrupting chemicals and effect biomarkers associating disease risks: where exposomics meets metabolomics. *Environ Int*. Jan 2022;158:106941. [FREE Full text] [doi: [10.1016/j.envint.2021.106941](https://doi.org/10.1016/j.envint.2021.106941)] [Medline: [34689039](https://pubmed.ncbi.nlm.nih.gov/34689039/)]
257. Blair A, Freeman LB. Epidemiologic studies in agricultural populations: observations and future directions. *J Agromedicine*. May 07, 2009;14(2):125-131. [FREE Full text] [doi: [10.1080/10599240902779436](https://doi.org/10.1080/10599240902779436)] [Medline: [19437268](https://pubmed.ncbi.nlm.nih.gov/19437268/)]
258. Carvalho JG, Augusto HC, Ferraz R, Delerue-Matos C, Fernandes VC. Micro(nano)plastic and related chemicals: emerging contaminants in environment, food and health impacts. *Toxics*. Oct 20, 2024;12(10):762. [FREE Full text] [doi: [10.3390/toxics12100762](https://doi.org/10.3390/toxics12100762)] [Medline: [39453182](https://pubmed.ncbi.nlm.nih.gov/39453182/)]
259. Clarke K, Manrique A, Sabo-Attwood T, Coker ES. A narrative review of occupational air pollution and respiratory health in farmworkers. *Int J Environ Res Public Health*. Apr 13, 2021;18(8):4097. [FREE Full text] [doi: [10.3390/ijerph18084097](https://doi.org/10.3390/ijerph18084097)] [Medline: [33924663](https://pubmed.ncbi.nlm.nih.gov/33924663/)]
260. Domingo NG, Balasubramanian S, Thakrar SK, Clark MA, Adams PJ, Marshall JD, et al. Air quality-related health damages of food. *Proc Natl Acad Sci U S A*. May 18, 2021;118(20):e2013637118. [FREE Full text] [doi: [10.1073/pnas.2013637118](https://doi.org/10.1073/pnas.2013637118)] [Medline: [33972419](https://pubmed.ncbi.nlm.nih.gov/33972419/)]
261. Hao Y, Sun H, Zeng X, Dong G, Kronzucker HJ, Min J, et al. Smallholder vegetable farming produces more soil microplastics pollution than large-scale farming. *Environ Pollut*. Jan 15, 2023;317:120805. [doi: [10.1016/j.envpol.2022.120805](https://doi.org/10.1016/j.envpol.2022.120805)] [Medline: [36470457](https://pubmed.ncbi.nlm.nih.gov/36470457/)]

262. Ng EL, Huerta Lwanga E, Eldridge SM, Johnston P, Hu H, Geissen V, et al. An overview of microplastic and nanoplastic pollution in agroecosystems. *Sci Total Environ*. Jun 15, 2018;627:1377-1388. [doi: [10.1016/j.scitotenv.2018.01.341](https://doi.org/10.1016/j.scitotenv.2018.01.341)] [Medline: [30857101](https://pubmed.ncbi.nlm.nih.gov/30857101/)]
263. van Kersen W, Bossers A, de Steenhuijsen Piters WA, de Rooij MM, Bonten M, Fluit AC, et al. Air pollution from livestock farms and the oropharyngeal microbiome of COPD patients and controls. *Environ Int*. Nov 2022;169:107497. [FREE Full text] [doi: [10.1016/j.envint.2022.107497](https://doi.org/10.1016/j.envint.2022.107497)] [Medline: [36088872](https://pubmed.ncbi.nlm.nih.gov/36088872/)]
264. Ganter M. Zoonotic risks from small ruminants. *Vet Microbiol*. Dec 14, 2015;181(1-2):53-65. [doi: [10.1016/j.vetmic.2015.07.015](https://doi.org/10.1016/j.vetmic.2015.07.015)] [Medline: [26275853](https://pubmed.ncbi.nlm.nih.gov/26275853/)]
265. Golomazou E, Mamedova S, Eslahi AV, Karanis P. Cryptosporidium and agriculture: a review. *Sci Total Environ*. Mar 15, 2024;916:170057. [doi: [10.1016/j.scitotenv.2024.170057](https://doi.org/10.1016/j.scitotenv.2024.170057)] [Medline: [38242460](https://pubmed.ncbi.nlm.nih.gov/38242460/)]
266. McDaniel CJ, Cardwell DM, Moeller Jr RB, Gray GC. Humans and cattle: a review of bovine zoonoses. *Vector Borne Zoonotic Dis*. Jan 2014;14(1):1-19. [FREE Full text] [doi: [10.1089/vbz.2012.1164](https://doi.org/10.1089/vbz.2012.1164)] [Medline: [24341911](https://pubmed.ncbi.nlm.nih.gov/24341911/)]
267. Arora K, Xu L, Bhagianadh D. Dementia and cognitive decline in older adulthood: are agricultural workers at greater risk? *J Gerontol B Psychol Sci Soc Sci*. Sep 13, 2021;76(8):1629-1643. [FREE Full text] [doi: [10.1093/geronb/gbab005](https://doi.org/10.1093/geronb/gbab005)] [Medline: [33406265](https://pubmed.ncbi.nlm.nih.gov/33406265/)]
268. Douphrate DI, Rosecrance JC, Reynolds SJ, Stallones L, Gilkey DP. Tractor-related injuries: an analysis of workers' compensation data. *J Agromedicine*. May 07, 2009;14(2):198-205. [doi: [10.1080/10599240902773215](https://doi.org/10.1080/10599240902773215)] [Medline: [19437278](https://pubmed.ncbi.nlm.nih.gov/19437278/)]
269. Dos Anjos Magri C, Garófallo Garcia R, Binotto E, Duarte da Silva Lima N, de Alencar Nääs I, Sgavioli S, et al. Occupational risk factors in health of broiler-farm workers: a systematic review. *Arch Environ Occup Health*. Oct 15, 2021;76(8):482-493. [doi: [10.1080/19338244.2020.1832036](https://doi.org/10.1080/19338244.2020.1832036)] [Medline: [33054688](https://pubmed.ncbi.nlm.nih.gov/33054688/)]
270. Barbosa Junior M, Sokulski CC, Salvador R, Pinheiro E, de Francisco AC, Trojan F. What kills the agricultural worker? A systematic review on suicide. *Rural Remote Health*. Aug 2021;21(3):6067. [FREE Full text] [doi: [10.22605/RRH6067](https://doi.org/10.22605/RRH6067)] [Medline: [34384223](https://pubmed.ncbi.nlm.nih.gov/34384223/)]
271. Garrett-Wright D, Malin C, Jones MS. Mental health in farming communities. *J Psychosoc Nurs Ment Health Serv*. Oct 2023;61(10):39-43. [doi: [10.3928/02793695-20230424-02](https://doi.org/10.3928/02793695-20230424-02)] [Medline: [37134283](https://pubmed.ncbi.nlm.nih.gov/37134283/)]
272. Lunner Kolstrup C, Kallioniemi M, Lundqvist P, Kymäläinen HR, Stallones L, Brumby S. International perspectives on psychosocial working conditions, mental health, and stress of dairy farm operators. *J Agromedicine*. Jul 03, 2013;18(3):244-255. [doi: [10.1080/1059924X.2013.796903](https://doi.org/10.1080/1059924X.2013.796903)] [Medline: [23844791](https://pubmed.ncbi.nlm.nih.gov/23844791/)]
273. Petit P, Gandon G, Dubuc M, Vuillerme N, Bonnetterre V. Agricultural activities and risk of treatment for depressive disorders among the entire French agricultural workforce: the TRACTOR project, a nationwide retrospective cohort study. *Lancet Reg Health Eur*. Aug 2023;31:100674. [FREE Full text] [doi: [10.1016/j.lanepe.2023.100674](https://doi.org/10.1016/j.lanepe.2023.100674)] [Medline: [37408876](https://pubmed.ncbi.nlm.nih.gov/37408876/)]
274. Daghigh Yazd S, Wheeler SA, Zuo A. Key risk factors affecting farmers' mental health: a systematic review. *Int J Environ Res Public Health*. Dec 02, 2019;16(23):4849. [FREE Full text] [doi: [10.3390/ijerph16234849](https://doi.org/10.3390/ijerph16234849)] [Medline: [31810320](https://pubmed.ncbi.nlm.nih.gov/31810320/)]
275. Nordby K, Andersen A, Irgens LM, Kristensen P. Indicators of mancozeb exposure in relation to thyroid cancer and neural tube defects in farmers' families. *Scand J Work Environ Health*. Apr 2005;31(2):89-96. [FREE Full text] [doi: [10.5271/sjweh.855](https://doi.org/10.5271/sjweh.855)] [Medline: [15864902](https://pubmed.ncbi.nlm.nih.gov/15864902/)]
276. Kauppinen T, Uksulainen S, Saalo A, Mäkinen I, Pukkala E. Use of the Finnish Information System on Occupational Exposure (FINJEM) in epidemiologic, surveillance, and other applications. *Ann Occup Hyg*. Apr 2014;58(3):380-396. [doi: [10.1093/annhyg/met074](https://doi.org/10.1093/annhyg/met074)] [Medline: [24401793](https://pubmed.ncbi.nlm.nih.gov/24401793/)]
277. Shim YK, Mlynarek SP, van Wijngaarden E. Parental exposure to pesticides and childhood brain cancer: U.S. Atlantic coast childhood brain cancer study. *Environ Health Perspect*. Jun 2009;117(6):1002-1006. [FREE Full text] [doi: [10.1289/ehp.0800209](https://doi.org/10.1289/ehp.0800209)] [Medline: [19590697](https://pubmed.ncbi.nlm.nih.gov/19590697/)]
278. Baldi I, Carles C, Blanc-Lapierre A, Fabbro-Peray P, Druet-Cabanac M, Boutet-Robinet E, et al. PESTIMAT Group. A French crop-exposure matrix for use in epidemiological studies on pesticides: PESTIMAT. *J Expo Sci Environ Epidemiol*. Jan 23, 2017;27(1):56-63. [doi: [10.1038/jes.2015.72](https://doi.org/10.1038/jes.2015.72)] [Medline: [26696463](https://pubmed.ncbi.nlm.nih.gov/26696463/)]
279. Oltramare C, Mediouni Z, Shoman Y, Hopf NB, Graczyk H, Berthet A. Determinants of pesticide exposure in occupational studies: a meta-analysis. *Toxics*. Jul 18, 2023;11(7):623. [FREE Full text] [doi: [10.3390/toxics11070623](https://doi.org/10.3390/toxics11070623)] [Medline: [37505588](https://pubmed.ncbi.nlm.nih.gov/37505588/)]
280. Burstyn I, Lavoué J, Van Tongeren M. Aggregation of exposure level and probability into a single metric in job-exposure matrices creates bias. *Ann Occup Hyg*. Nov 2012;56(9):1038-1050. [doi: [10.1093/annhyg/mes031](https://doi.org/10.1093/annhyg/mes031)] [Medline: [22986426](https://pubmed.ncbi.nlm.nih.gov/22986426/)]
281. Fadel M, Evanoff BA, Andersen JH, d'Errico A, Dale AM, Leclerc A, et al. Not just a research method: if used with caution, can job-exposure matrices be a useful tool in the practice of occupational medicine and public health? *Scand J Work Environ Health*. Sep 01, 2020;46(5):552-553. [FREE Full text] [doi: [10.5271/sjweh.3900](https://doi.org/10.5271/sjweh.3900)] [Medline: [32367143](https://pubmed.ncbi.nlm.nih.gov/32367143/)]
282. Peters S, Vienneau D, Sampri A, Turner MC, Castaño-Vinyals G, Bugge M, et al. Occupational exposure assessment tools in Europe: a comprehensive inventory overview. *Ann Work Expo Health*. Jun 06, 2022;66(5):671-686. [FREE Full text] [doi: [10.1093/annweh/wxab110](https://doi.org/10.1093/annweh/wxab110)] [Medline: [34935027](https://pubmed.ncbi.nlm.nih.gov/34935027/)]
283. Petit P, Gondard E, Gandon G, Moreaud O, Sauvée M, Bonnetterre V. Agricultural activities and risk of Alzheimer's disease: the TRACTOR project, a nationwide retrospective cohort study. *Eur J Epidemiol*. Mar 10, 2024;39(3):271-287. [FREE Full text] [doi: [10.1007/s10654-023-01079-0](https://doi.org/10.1007/s10654-023-01079-0)] [Medline: [38195954](https://pubmed.ncbi.nlm.nih.gov/38195954/)]

284. List of classifications. International Agency for Research on Cancer. URL: <https://monographs.iarc.who.int/list-of-classifications> [accessed 2024-10-30]
285. Desai B, Sahni S, Jordan H, Sahni R, Reinbeau R, Nguyen A, et al. Risk of cardiovascular disease in male farmers over the age of 45: a review of literature. *Cureus*. May 2022;14(5):e24642. [FREE Full text] [doi: [10.7759/cureus.24642](https://doi.org/10.7759/cureus.24642)] [Medline: [35663675](https://pubmed.ncbi.nlm.nih.gov/35663675/)]
286. Kim H, Jung W, Jung S, Cho S, Jung I, Song H, et al. Is farming a risk occupation for cardio-cerebrovascular diseases? A scoping review on cardio-cerebrovascular disease risk in farmers. *J Prev Med Public Health*. Aug 23, 2024;57(6):521-529. [FREE Full text] [doi: [10.3961/jpmph.24.302](https://doi.org/10.3961/jpmph.24.302)] [Medline: [39438011](https://pubmed.ncbi.nlm.nih.gov/39438011/)]
287. Hulkkonen S, Shiri R, Auvinen J, Miettunen J, Karppinen J, Ryhänen J. Risk factors of hospitalization for carpal tunnel syndrome among the general working population. *Scand J Work Environ Health*. Jan 01, 2020;46(1):43-49. [FREE Full text] [doi: [10.5271/sjweh.3835](https://doi.org/10.5271/sjweh.3835)] [Medline: [31168633](https://pubmed.ncbi.nlm.nih.gov/31168633/)]
288. Kim J, Youn K, Park J. Risk factors for musculoskeletal disorders in Korean farmers: survey on occupational diseases in 2020 and 2022. *Healthcare (Basel)*. Oct 11, 2024;12(20):2026. [FREE Full text] [doi: [10.3390/healthcare12202026](https://doi.org/10.3390/healthcare12202026)] [Medline: [39451441](https://pubmed.ncbi.nlm.nih.gov/39451441/)]
289. Hemminki K, Niemi ML, Saloniemi I, Vainio H, Hemminki E. Spontaneous abortions by occupation and social class in Finland. *Int J Epidemiol*. Jun 1980;9(2):149-153. [doi: [10.1093/ije/9.2.149](https://doi.org/10.1093/ije/9.2.149)] [Medline: [7409966](https://pubmed.ncbi.nlm.nih.gov/7409966/)]
290. Strand K, Wergeland E, Bjerkedal T. Fertility patterns according to occupational grouping in Norway, 1989. *Scand J Soc Med*. Mar 01, 1996;24(1):50-54. [doi: [10.1177/140349489602400108](https://doi.org/10.1177/140349489602400108)] [Medline: [8740876](https://pubmed.ncbi.nlm.nih.gov/8740876/)]
291. Liu Y, Huang L, Hu F, Zhang X. Investigating frailty, polypharmacy, malnutrition, chronic conditions, and quality of life in older adults: large population-based study. *JMIR Public Health Surveill*. Oct 11, 2024;10:e50617. [FREE Full text] [doi: [10.2196/50617](https://doi.org/10.2196/50617)] [Medline: [39145920](https://pubmed.ncbi.nlm.nih.gov/39145920/)]
292. Lie A, Skogstad M, Johannessen HA, Tynes T, Mehlum IS, Nordby K, et al. Occupational noise exposure and hearing: a systematic review. *Int Arch Occup Environ Health*. Apr 7, 2016;89(3):351-372. [FREE Full text] [doi: [10.1007/s00420-015-1083-5](https://doi.org/10.1007/s00420-015-1083-5)] [Medline: [26249711](https://pubmed.ncbi.nlm.nih.gov/26249711/)]
293. Kraemer JG, Aebi S, Oppliger A, Hilty M. The Indoor-Air Microbiota of pig farms drives the composition of the pig farmers' nasal microbiota in a season-dependent and farm-specific manner. *Appl Environ Microbiol*. May 01, 2019;85(9):e03038. [FREE Full text] [doi: [10.1128/AEM.03038-18](https://doi.org/10.1128/AEM.03038-18)] [Medline: [30824439](https://pubmed.ncbi.nlm.nih.gov/30824439/)]
294. Mucci N, Tommasi E, Chiarelli A, Lulli LG, Traversini V, Galea RP, et al. WORKbiota: a systematic review about the effects of occupational exposure on microbiota and workers' health. *Int J Environ Res Public Health*. Jan 18, 2022;19(3):1043. [FREE Full text] [doi: [10.3390/ijerph19031043](https://doi.org/10.3390/ijerph19031043)] [Medline: [35162072](https://pubmed.ncbi.nlm.nih.gov/35162072/)]
295. Shukla SK, Ye Z, Sandberg S, Reyes I, Fritsche TR, Keifer M. The nasal microbiota of dairy farmers is more complex than oral microbiota, reflects occupational exposure, and provides competition for staphylococci. *PLoS One*. Aug 29, 2017;12(8):e0183898. [FREE Full text] [doi: [10.1371/journal.pone.0183898](https://doi.org/10.1371/journal.pone.0183898)] [Medline: [28850578](https://pubmed.ncbi.nlm.nih.gov/28850578/)]
296. Sun J, Liao XP, D'Souza AW, Boolchandani M, Li S, Cheng K, et al. Environmental remodeling of human gut microbiota and antibiotic resistome in livestock farms. *Nat Commun*. Mar 18, 2020;11(1):1427. [FREE Full text] [doi: [10.1038/s41467-020-15222-y](https://doi.org/10.1038/s41467-020-15222-y)] [Medline: [32188862](https://pubmed.ncbi.nlm.nih.gov/32188862/)]
297. Vilela C, Araújo B, Soares-Guedes C, Caridade-Silva R, Martins-Macedo J, Teixeira C, et al. From the gut to the brain: is microbiota a new paradigm in Parkinson's disease treatment? *Cells*. Apr 30, 2024;13(9):770. [FREE Full text] [doi: [10.3390/cells13090770](https://doi.org/10.3390/cells13090770)] [Medline: [38727306](https://pubmed.ncbi.nlm.nih.gov/38727306/)]
298. Gois MF, Fernández-Pato A, Huss A, Gacesa R, Wijmenga C, Weersma RK, et al. Impact of occupational pesticide exposure on the human gut microbiome. *Front Microbiol*. Aug 10, 2023;14:1223120. [FREE Full text] [doi: [10.3389/fmicb.2023.1223120](https://doi.org/10.3389/fmicb.2023.1223120)] [Medline: [37637104](https://pubmed.ncbi.nlm.nih.gov/37637104/)]
299. Christen P, Schnell R. Thirty-three myths and misconceptions about population data: from data capture and processing to linkage. *Int J Popul Data Sci*. Jan 31, 2023;8(1):2115. [FREE Full text] [doi: [10.23889/ijpds.v8i1.2115](https://doi.org/10.23889/ijpds.v8i1.2115)] [Medline: [37636835](https://pubmed.ncbi.nlm.nih.gov/37636835/)]
300. Dattani N, Hardelid P, Davey J, Gilbert R. Accessing electronic administrative health data for research takes time. *Arch Dis Child*. May 28, 2013;98(5):391-392. [doi: [10.1136/archdischild-2013-303730](https://doi.org/10.1136/archdischild-2013-303730)] [Medline: [23448863](https://pubmed.ncbi.nlm.nih.gov/23448863/)]
301. Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible sources of bias in primary care electronic health record data use and reuse. *J Med Internet Res*. May 29, 2018;20(5):e185. [FREE Full text] [doi: [10.2196/jmir.9134](https://doi.org/10.2196/jmir.9134)] [Medline: [29844010](https://pubmed.ncbi.nlm.nih.gov/29844010/)]
302. Bousquet P, Caillet P, Coeuret-Pellicer M, Goulard H, Kudjawa Y, Le Bihan C, et al. [Using cancer case identification algorithms in medico-administrative databases: literature review and first results from the REDSIAM Tumors group based on breast, colon, and lung cancer]. *Rev Epidemiol Sante Publique*. Oct 2017;65 Suppl 4:S236-S242. [doi: [10.1016/j.respe.2017.04.057](https://doi.org/10.1016/j.respe.2017.04.057)] [Medline: [28576379](https://pubmed.ncbi.nlm.nih.gov/28576379/)]
303. Goldberg M, Carton M, Doussin A, Fagot-Campagna A, Heyndrickx E, Lemaitre M, et al. [The REDSIAM network]. *Rev Epidemiol Sante Publique*. Oct 2017;65 Suppl 4:S144-S148. [doi: [10.1016/j.respe.2017.06.001](https://doi.org/10.1016/j.respe.2017.06.001)] [Medline: [28844426](https://pubmed.ncbi.nlm.nih.gov/28844426/)]
304. Lipscombe LL, Hwee J, Webster L, Shah BR, Booth GL, Tu K. Identifying diabetes cases from administrative data: a population-based validation study. *BMC Health Serv Res*. May 02, 2018;18(1):316. [FREE Full text] [doi: [10.1186/s12913-018-3148-0](https://doi.org/10.1186/s12913-018-3148-0)] [Medline: [29720153](https://pubmed.ncbi.nlm.nih.gov/29720153/)]

305. Caputo M, Pecere A, Sarro A, Mele C, Ucciero A, Pagano L, et al. Incidence and prevalence of hyperthyroidism: a population-based study in the Piedmont Region, Italy. *Endocrine*. Jul 13, 2020;69(1):107-112. [doi: [10.1007/s12020-020-02222-7](https://doi.org/10.1007/s12020-020-02222-7)] [Medline: [32056093](https://pubmed.ncbi.nlm.nih.gov/32056093/)]
306. Habib RR, Hojeij S, Elzein K. Gender in occupational health research of farmworkers: a systematic review. *Am J Ind Med*. Dec 16, 2014;57(12):1344-1367. [FREE Full text] [doi: [10.1002/ajim.22375](https://doi.org/10.1002/ajim.22375)] [Medline: [25227724](https://pubmed.ncbi.nlm.nih.gov/25227724/)]
307. Biswas A, Harbin S, Irvin E, Johnston H, Begum M, Tiong M, et al. Sex and gender differences in occupational hazard exposures: a scoping review of the recent literature. *Curr Environ Health Rep*. Dec 27, 2021;8(4):267-280. [FREE Full text] [doi: [10.1007/s40572-021-00330-8](https://doi.org/10.1007/s40572-021-00330-8)] [Medline: [34839446](https://pubmed.ncbi.nlm.nih.gov/34839446/)]
308. Biswas A, Tiong M, Irvin E, Zhai G, Sinkins M, Johnston H, et al. Gender and sex differences in occupation-specific infectious diseases: a systematic review. *Occup Environ Med*. Aug 29, 2024;81(8):425-432. [FREE Full text] [doi: [10.1136/oemed-2024-109451](https://doi.org/10.1136/oemed-2024-109451)] [Medline: [39168602](https://pubmed.ncbi.nlm.nih.gov/39168602/)]
309. Comfort N, Re DB. Sex-specific neurotoxic effects of organophosphate pesticides across the life course. *Curr Environ Health Rep*. Dec 23, 2017;4(4):392-404. [FREE Full text] [doi: [10.1007/s40572-017-0171-y](https://doi.org/10.1007/s40572-017-0171-y)] [Medline: [29063415](https://pubmed.ncbi.nlm.nih.gov/29063415/)]
310. Wang W, Jin J, He R, Gong H. Gender differences in pesticide use knowledge, risk awareness and practices in Chinese farmers. *Sci Total Environ*. Jul 15, 2017;590-591:22-28. [doi: [10.1016/j.scitotenv.2017.03.053](https://doi.org/10.1016/j.scitotenv.2017.03.053)] [Medline: [28285132](https://pubmed.ncbi.nlm.nih.gov/28285132/)]
311. Xu L, Xu X, Wu X, Kuang H, Xu C. Sex-dependent environmental health risk analysis of flupyradifurone. *Environ Sci Technol*. Feb 01, 2022;56(3):1841-1853. [doi: [10.1021/acs.est.1c07726](https://doi.org/10.1021/acs.est.1c07726)] [Medline: [35041393](https://pubmed.ncbi.nlm.nih.gov/35041393/)]
312. Ketchen Jr DJ, Roccapriore AY, Connelly BL. Using old data: when is it appropriate? *J Manag*. Jun 04, 2023;49(8):2541-2548. [doi: [10.1177/01492063231177785](https://doi.org/10.1177/01492063231177785)]
313. Vallot C. Epigenomic tumor evolution under the spotlight: the promises of single-cell approaches. *C R Biol*. May 11, 2022;345(1):11-16. [FREE Full text] [doi: [10.5802/crbior.75](https://doi.org/10.5802/crbior.75)] [Medline: [35787617](https://pubmed.ncbi.nlm.nih.gov/35787617/)]
314. Horgan D, Hajduch M, Vrana M, Soderberg J, Hughes N, Omar MI, et al. European health data space-an opportunity now to grasp the future of data-driven healthcare. *Healthcare (Basel)*. Aug 26, 2022;10(9):1629. [FREE Full text] [doi: [10.3390/healthcare10091629](https://doi.org/10.3390/healthcare10091629)] [Medline: [36141241](https://pubmed.ncbi.nlm.nih.gov/36141241/)]
315. Martani A, Geneviève LD, Wangmo T, Maurer J, Cramer K, Erard F, et al. Sensing the (digital) pulse. Future steps for improving the secondary use of data for research in Switzerland. *Digit Health*. Apr 20, 2023;9:20552076231169826. [FREE Full text] [doi: [10.1177/20552076231169826](https://doi.org/10.1177/20552076231169826)] [Medline: [37113255](https://pubmed.ncbi.nlm.nih.gov/37113255/)]
316. Raab R, Küderle A, Zakreuskaya A, Stern AD, Klucken J, Kaissis G, et al. Federated electronic health records for the European Health Data Space. *Lancet Digit Health*. Nov 2023;5(11):e840-e847. [FREE Full text] [doi: [10.1016/S2589-7500\(23\)00156-5](https://doi.org/10.1016/S2589-7500(23)00156-5)] [Medline: [37741765](https://pubmed.ncbi.nlm.nih.gov/37741765/)]
317. Pampel H, Weisweiler NL, Strecker D, Witt M, Vierkant P, Elger K, et al. re3data - indexing the global research data repository landscape since 2012. *Sci Data*. Aug 29, 2023;10(1):571. [FREE Full text] [doi: [10.1038/s41597-023-02462-y](https://doi.org/10.1038/s41597-023-02462-y)] [Medline: [37644024](https://pubmed.ncbi.nlm.nih.gov/37644024/)]
318. Suhr B, Dungal J, Stocker A. Search, reuse and sharing of research data in materials science and engineering-a qualitative interview study. *PLoS One*. Sep 15, 2020;15(9):e0239216. [FREE Full text] [doi: [10.1371/journal.pone.0239216](https://doi.org/10.1371/journal.pone.0239216)] [Medline: [32931508](https://pubmed.ncbi.nlm.nih.gov/32931508/)]
319. Noaen M, Doiron C, Syer J, Brook J. Advancing population health through open environmental data platforms. *Curr Top Behav Neurosci*. 2024;68:297-323. [doi: [10.1007/7854_2024_512](https://doi.org/10.1007/7854_2024_512)] [Medline: [39112811](https://pubmed.ncbi.nlm.nih.gov/39112811/)]
320. Dul J, van Raaij E, Caputo A. Advancing scientific inquiry through data reuse: necessary condition analysis with archival data. *Strateg Change*. Nov 02, 2023;33(1):35-40. [doi: [10.1002/jsc.2562](https://doi.org/10.1002/jsc.2562)]
321. Catalogue des bases de données individuelles en santé en France. INSERM. 2020. URL: <https://epidemiologie-france.aviesan.fr/> [accessed 2024-10-30]
322. Plateforme ouverte des données publiques françaises. République Française. URL: <https://www.data.gouv.fr/fr/> [accessed 2024-10-30]
323. Dataset search. Google. URL: <https://datasetsearch.research.google.com/> [accessed 2024-10-30]
324. Kogevinas M, Schlünssen V, Mehlum IS, Turner MC. The OMEGA-NET international inventory of occupational cohorts. *Ann Work Expo Health*. Jul 01, 2020;64(6):565-568. [FREE Full text] [doi: [10.1093/annweh/wxaa039](https://doi.org/10.1093/annweh/wxaa039)] [Medline: [32556221](https://pubmed.ncbi.nlm.nih.gov/32556221/)]
325. Peters S. Although a valuable method in occupational epidemiology, job-exposure -matrices are no magic fix. *Scand J Work Environ Health*. May 01, 2020;46(3):231-234. [FREE Full text] [doi: [10.5271/sjweh.3894](https://doi.org/10.5271/sjweh.3894)] [Medline: [32356897](https://pubmed.ncbi.nlm.nih.gov/32356897/)]
326. Nittas V, Puhan MA, von Wyl V. Toward a working definition of eCohort studies in health research: narrative literature review. *JMIR Public Health Surveill*. Jan 21, 2021;7(1):e24588. [FREE Full text] [doi: [10.2196/24588](https://doi.org/10.2196/24588)] [Medline: [33475521](https://pubmed.ncbi.nlm.nih.gov/33475521/)]
327. Singh AV, Bansod G, Mahajan M, Dietrich P, Singh SP, Rav K, et al. Digital transformation in toxicology: improving communication and efficiency in risk assessment. *ACS Omega*. Jun 20, 2023;8(24):21377-21390. [doi: [10.1021/acsomega.3c00596](https://doi.org/10.1021/acsomega.3c00596)] [Medline: [37360489](https://pubmed.ncbi.nlm.nih.gov/37360489/)]
328. Soussi A, Zero E, Sacile R, Trincherro D, Fossa M. Smart sensors and smart data for precision agriculture: a review. *Sensors (Basel)*. Apr 21, 2024;24(8):2647. [FREE Full text] [doi: [10.3390/s24082647](https://doi.org/10.3390/s24082647)] [Medline: [38676264](https://pubmed.ncbi.nlm.nih.gov/38676264/)]
329. Bukhtiyarova O, Abderrazak A, Chiu Y, Sparano S, Simard M, Sirois C. Major areas of interest of artificial intelligence research applied to health care administrative data: a scoping review. *Front Pharmacol*. Jul 18, 2022;13:944516. [FREE Full text] [doi: [10.3389/fphar.2022.944516](https://doi.org/10.3389/fphar.2022.944516)] [Medline: [35924057](https://pubmed.ncbi.nlm.nih.gov/35924057/)]

330. Caruana A, Bandara M, Musial K, Catchpoole D, Kennedy PJ. Machine learning for administrative health records: a systematic review of techniques and applications. *Artif Intell Med*. Oct 2023;144:102642. [doi: [10.1016/j.artmed.2023.102642](https://doi.org/10.1016/j.artmed.2023.102642)] [Medline: [37783537](https://pubmed.ncbi.nlm.nih.gov/37783537/)]
331. Leal Neto O, Von Wyl V. Digital transformation of public health for noncommunicable diseases: narrative viewpoint of challenges and opportunities. *JMIR Public Health Surveill*. Jan 25, 2024;10:e49575. [FREE Full text] [doi: [10.2196/49575](https://doi.org/10.2196/49575)] [Medline: [38271097](https://pubmed.ncbi.nlm.nih.gov/38271097/)]
332. Wang J, Kharrat FG, Gariépy G, Gagné C, Pelletier J, Massamba VK, et al. Predicting the population risk of suicide using routinely collected health administrative data in Quebec, Canada: model-based synthetic estimation study. *JMIR Public Health Surveill*. Jun 28, 2024;10:e52773. [FREE Full text] [doi: [10.2196/52773](https://doi.org/10.2196/52773)] [Medline: [38941610](https://pubmed.ncbi.nlm.nih.gov/38941610/)]
333. Ben-Ari T, Boé J, Ciais P, Lecerc R, Van der Velde M, Makowski D. Causes and implications of the unforeseen 2016 extreme yield loss in the breadbasket of France. *Nat Commun*. Apr 24, 2018;9(1):1627. [FREE Full text] [doi: [10.1038/s41467-018-04087-x](https://doi.org/10.1038/s41467-018-04087-x)] [Medline: [29691405](https://pubmed.ncbi.nlm.nih.gov/29691405/)]
334. Bhawra J, Elsahli N, Patel J. Applying digital technology to understand human experiences of climate change impacts on food security and mental health: scoping review. *JMIR Public Health Surveill*. Jul 23, 2024;10:e54064. [FREE Full text] [doi: [10.2196/54064](https://doi.org/10.2196/54064)] [Medline: [39042453](https://pubmed.ncbi.nlm.nih.gov/39042453/)]
335. Comi M, Becot F, Bendixsen C. Automation, climate change, and the future of farm work: cross-disciplinary lessons for studying dynamic changes in agricultural health and safety. *Int J Environ Res Public Health*. Mar 08, 2023;20(6):4778. [FREE Full text] [doi: [10.3390/ijerph20064778](https://doi.org/10.3390/ijerph20064778)] [Medline: [36981685](https://pubmed.ncbi.nlm.nih.gov/36981685/)]
336. Ebi KL, Vanos J, Baldwin JW, Bell JE, Hondula DM, Errett NA, et al. Extreme weather and climate change: population health and health system implications. *Annu Rev Public Health*. Apr 01, 2021;42:293-315. [FREE Full text] [doi: [10.1146/annurev-publhealth-012420-105026](https://doi.org/10.1146/annurev-publhealth-012420-105026)] [Medline: [33406378](https://pubmed.ncbi.nlm.nih.gov/33406378/)]
337. El Khayat M, Halwani DA, Hneiny L, Alameddine I, Haidar MA, Habib RR. Impacts of climate change and heat stress on farmworkers' health: a scoping review. *Front Public Health*. Feb 8, 2022;10:782811. [FREE Full text] [doi: [10.3389/fpubh.2022.782811](https://doi.org/10.3389/fpubh.2022.782811)] [Medline: [35211437](https://pubmed.ncbi.nlm.nih.gov/35211437/)]
338. Shahzad A, Ullah S, Dar AA, Sardar MF, Mehmood T, Tufail MA, et al. Nexus on climate change: agriculture and possible solution to cope future climate change stresses. *Environ Sci Pollut Res*. Jan 29, 2021;28(12):14211-14232. [doi: [10.1007/S11356-021-12649-8](https://doi.org/10.1007/S11356-021-12649-8)]
339. Gibb K, Beckman S, Vergara XP, Heinzerling A, Harrison R. Extreme heat and occupational health risks. *Annu Rev Public Health*. May 2024;45(1):315-335. [FREE Full text] [doi: [10.1146/annurev-publhealth-060222-034715](https://doi.org/10.1146/annurev-publhealth-060222-034715)] [Medline: [38166501](https://pubmed.ncbi.nlm.nih.gov/38166501/)]
340. Lefèvre-Arbogast S, Chaker J, Mercier F, Barouki R, Coumoul X, Miller GW, et al. Assessing the contribution of the chemical exposome to neurodegenerative disease. *Nat Neurosci*. May 29, 2024;27(5):812-821. [FREE Full text] [doi: [10.1038/s41593-024-01627-1](https://doi.org/10.1038/s41593-024-01627-1)] [Medline: [38684891](https://pubmed.ncbi.nlm.nih.gov/38684891/)]
341. Shan L, Heusinkveld HJ, Paul KC, Hughes S, Darweesh SK, Bloem BR, et al. Towards improved screening of toxins for Parkinson's risk. *NPJ Parkinsons Dis*. Dec 19, 2023;9(1):169. [FREE Full text] [doi: [10.1038/s41531-023-00615-9](https://doi.org/10.1038/s41531-023-00615-9)] [Medline: [38114496](https://pubmed.ncbi.nlm.nih.gov/38114496/)]
342. Abdelzaher H, Tawfik SM, Nour A, Abdelkader S, Elbalkiny ST, Abdelkader M, et al. Climate change, human health, and the exposome: utilizing OMIC technologies to navigate an era of uncertainty. *Front Public Health*. Sep 21, 2022;10:973000. [FREE Full text] [doi: [10.3389/fpubh.2022.973000](https://doi.org/10.3389/fpubh.2022.973000)] [Medline: [36211706](https://pubmed.ncbi.nlm.nih.gov/36211706/)]
343. DeBord DG, Carreón T, Lentz TJ, Middendorf PJ, Hoover MD, Schulte PA. Use of the "exposome" in the practice of epidemiology: a primer on -Omic technologies. *Am J Epidemiol*. Aug 15, 2016;184(4):302-314. [FREE Full text] [doi: [10.1093/aje/kwv325](https://doi.org/10.1093/aje/kwv325)] [Medline: [27519539](https://pubmed.ncbi.nlm.nih.gov/27519539/)]
344. Waffenschmidt S, Knelangen M, Sieben W, Bühn S, Pieper D. Single screening versus conventional double screening for study selection in systematic reviews: a methodological systematic review. *BMC Med Res Methodol*. Jun 28, 2019;19(1):132. [FREE Full text] [doi: [10.1186/s12874-019-0782-0](https://doi.org/10.1186/s12874-019-0782-0)] [Medline: [31253092](https://pubmed.ncbi.nlm.nih.gov/31253092/)]
345. Waffenschmidt S, Sieben W, Jakubeit T, Knelangen M, Overesch I, Bühn S, et al. Increasing the efficiency of study selection for systematic reviews using prioritization tools and a single-screening approach. *Syst Rev*. Sep 14, 2023;12(1):161. [FREE Full text] [doi: [10.1186/s13643-023-02334-x](https://doi.org/10.1186/s13643-023-02334-x)] [Medline: [37705060](https://pubmed.ncbi.nlm.nih.gov/37705060/)]

Abbreviations

- AGR:** annual growth rate
AGRICAN: Agriculture and Cancer
AGRICOH: agricultural cohort
AHD: administrative health database
AHS: Agricultural Health Study
AI: artificial intelligence
AMI: Aging Multidisciplinary Investigation
ANR: French National Research Agency (Agence Nationale de la Recherche in French)
AS: association strength

BIBLIO: preliminary guideline for reporting bibliometric reviews of the biomedical literature
COPD: chronic obstructive pulmonary disease
EHR: electronic health record
EMR: electronic medical record
FAIR: findable, accessible, interoperable, and reusable
FERMA: risk factors of the rural environment and the allergic and respiratory disease
GDP: gross domestic product
JEM: job-exposure matrix
MIAI: Multidisciplinary Institute in Artificial Intelligence
MSA: Mutualité Sociale Agricole
NOCCA: Nordic Occupational Cancer Study
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
TRACTOR: Tracking and Monitoring Occupational Risks in Agriculture

Edited by A Mavragani; submitted 05.06.24; peer-reviewed by G Wu, F Clerc; comments to author 30.09.24; revised version received 08.10.24; accepted 07.11.24; published 09.01.25

Please cite as:

Petit P, Vuillerme N

Leveraging Administrative Health Databases to Address Health Challenges in Farming Populations: Scoping Review and Bibliometric Analysis (1975-2024)

JMIR Public Health Surveill 2025;11:e62939

URL: <https://publichealth.jmir.org/2025/1/e62939>

doi: [10.2196/62939](https://doi.org/10.2196/62939)

PMID:

©Pascal Petit, Nicolas Vuillerme. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 09.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.

Multimedia Appendices

Title: Leveraging administrative health databases to address health challenges in farming populations: a scoping review and bibliometric analysis (1975-2024)

Authors

Pascal Petit^{1*}; Nicolas Vuillerme^{1,2}

¹ Univ. Grenoble Alpes, AGEIS, 38000 Grenoble, France

² Institut Universitaire de France, Paris, France

Correspondence (present address)

* Corresponding author at: Pascal Petit. Laboratoire AGEIS – Université Grenoble Alpes. Bureau 315. Bâtiment Jean Roget. UFR de Médecine. Domaine de La Merci. 38706 La Tronche Cedex, France.

E-mail address: pascal.petit@univ-grenoble-alpes.fr (P. Petit). Telephone: + 33 4 76 63 71 04.

Table of content

Table S1. Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) Checklist.	3
Table S2. Criteria used for the publication selection.	4
Table S3. Bibliometric index metrics.	5
Table S4. The BIBLIO checklist for reporting the bibliometric reviews of the biomedical literature.	6
Figure S1. Publication type.	7
Figure S2. Scientific productivity in terms of publications per period of time.	8
Figure S3. Historical direct citation network.	9
Figure S4. Most prolific authors' country.	11
Table S5. Top 25 of the most cited publications.	12
Table S6. Other bibliometric analysis information.	13
Multimedia appendix 2. Publications included and analyzed.	20
Figure S5. Number of publications per country.	21
Figure S6. Main goal addressed by study.	22
Figure S7. Number of publications per cohort.	23
Figure S8. Type of digital data used.	24
Figure S9. Active data used.	25
Figure S10. Most frequent farming exposure proxy.	26
Figure S11. Most frequent health events studied.	27
Figure S12. Most frequent cancer studied.	28
Figure S13. Most frequent neurodegenerative disease studied.	29
Figure S14. Most frequent mental health issue studied.	30
Table S7. Top 50 of the most frequent keywords.	31
Figure S15. Top 5 of the most frequent keywords – growth trends.	33
Figure S16. Top ten of the most popular keywords by total number of years of appearance.	34
Figure S17. Top 10 of the most frequent keywords by time period.	35
Multimedia appendix 3. Co-occurrence between exposome-related keywords and health event-related keywords.	36
Figure S18. Keyword co-occurrence network between potential risk factor and mental health disorder keywords.	37
Figure S19. Chord diagram of keyword co-occurrence between potential risk factor and neurodegenerative disease keywords.	38
Figure S20. Keyword co-occurrence network between potential risk factor and neurodegenerative disease keywords.	39
Figure S21. Chord diagram of keyword co-occurrence between potential risk factor and autoimmune disease keywords.	40
Figure S22. Keyword co-occurrence network between potential risk factor and autoimmune disease keywords.	41
Figure S23. Chord diagram of keyword co-occurrence between potential risk factor and breast and genital organ cancer keywords.	42

60 **Figure S24.** Keyword co-occurrence network between potential risk factor and breast and genital organ cancer
61 keywords. 43
62 **Figure S25.** Chord diagram of keyword co-occurrence between potential risk factor and digestive organ, lip, oral
63 cavity and pharynx cancer keywords. 44
64 **Figure S26.** Keyword co-occurrence network between potential risk factor and digestive organ, lip, oral cavity
65 and pharynx cancer keywords. 45
66 **Figure S27.** Chord diagram of keyword co-occurrence between potential risk factor and lymphohematopoietic and
67 mesenchymal cancer keywords. 46
68 **Figure S28.** Keyword co-occurrence network between potential risk factor and lymphohematopoietic and
69 mesenchymal cancer keywords. 47
70 **Figure S29.** Chord diagram of keyword co-occurrence between potential risk factor and brain, skin, ocular and
71 endocrine gland cancer keywords. 48
72 **Figure S30.** Keyword co-occurrence network between potential risk factor and brain, skin, ocular and endocrine
73 gland cancer keywords. 49
74 **Figure S31.** Chord diagram of keyword co-occurrence between potential risk factor and respiratory and urinary
75 tract cancer keywords. 50
76 **Figure S32.** Keyword co-occurrence network between potential risk factor and respiratory and urinary tract cancer
77 keywords. 51
78 **Figure S33.** Chord diagram of keyword co-occurrence between potential risk factor and infectious disease
79 keywords. 52
80 **Figure S34.** Keyword co-occurrence network between potential risk factor and infectious disease keywords. .. 53
81 **Figure S35.** Chord diagram of keyword co-occurrence between potential risk factor and respiratory disorder
82 keywords. 54
83 **Figure S36.** Keyword co-occurrence network between potential risk factor and respiratory disorder keywords. 55
84 **Figure S37.** Chord diagram of keyword co-occurrence between potential risk factor and reproductive disorder
85 keywords. 56
86 **Figure S38.** Keyword co-occurrence network between potential risk factor and reproductive disorder keywords.
87 57
88 **Figure S39.** Chord diagram of keyword co-occurrence between potential risk factor and injury and work-related
89 disease keywords. 58
90 **Figure S40.** Keyword co-occurrence network between potential risk factor and injury and work-related disease
91 keywords. 59
92 **Figure S41.** Chord diagram of keyword co-occurrence between potential risk factor and other health event
93 keywords. 60
94 **Figure S42.** Keyword co-occurrence network between potential risk factor and other health event keywords. .. 61
95 **Figure S43.** Keyword co-occurrence network between potential risk factor and mortality keywords. 62
96 **Abbreviations** 63
97 **References** 64
98
99

100 **Table S1.** Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews
 101 (PRISMA-ScR) Checklist.
 102

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a scoping review.	1
ABSTRACT			
Structured summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	2, 3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	4-7
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	7
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	8
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	8-10
Information sources*	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	8-10
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	9
Selection of sources of evidence†	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	8-10
Data charting process‡	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	10-14
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	10-14
Critical appraisal of individual sources of evidence§	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	NA
Synthesis of results	13	Describe the methods of handling and summarizing the data that were charted.	10-14
RESULTS			
Selection of sources of evidence	14	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	14-17, supplementary materials
Characteristics of sources of evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	14-26, supplementary materials
Critical appraisal within sources of evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	NA
Results of individual sources of evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	14-26, supplementary materials
Synthesis of results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	14-39, supplementary materials
DISCUSSION			
Summary of evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	40-47
Limitations	20	Discuss the limitations of the scoping review process.	48, 49
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	49-51
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	52

103 JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping
 104 Reviews.

105 * Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.

106 † A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative
 107 research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused
 108 with *information sources* (see first footnote).

109 ‡ The frameworks by Arksey and O'Malley (6) and Levac and colleagues (7) and the JBI guidance (4, 5) refer to the process of data extraction
 110 in a scoping review as data charting.

111 § The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision.
 112 This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and
 113 acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion,
 114 and policy document).

115 Peters MDJ, Godfrey C, McInerney P, BaldiniSoares C, Khalil H, Parker D. Scoping reviews. In: Aromataris E, Munn Z, eds. Joanna Briggs Institute Reviewer's
 116 Manual. Adelaide, Australia: Joanna Briggs Inst; 2017.

117 Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005;8:19-32. doi: 10.1080/1364557032000119616

118 Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. *Int J Evid Based Healthc.*
 119 2015;13(3):141-146. doi:10.1097/XEB.0000000000000050

120 Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci.* 2010;5:69. Published 2010 Sep 20. doi:10.1186/1748-5908-5-
 121 69.

122 *From:* Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR):
 123 Checklist and Explanation. *Ann Intern Med.* 2018;169:467-473. doi: 10.7326/M18-0850.

124 **Table S2.** Criteria used for the publication selection.
 125

Question	Description	Answer	
		no	yes/can't tell
Stage 1: Record language			
Q_{11}	Is the record written in English or French?	0	1
$S_1 = Q_{11}$; Record eligible for stage 2 if $S_1 > 0$			
Stage 2: Screening record title			
Q_{21}	Does the title mention terms related to administrative health data?	0	1
Q_{22}	Does the title mention terms related to humans?	0	1
Q_{23}	Does the title mention terms related to farming?	0	1
Q_{24}	Does the title mention terms related to public health or epidemiological research?	0	1
$S_2 = Q_{21} \times Q_{22} \times Q_{23} \times Q_{24}$; Record eligible for stage 3 if $S_2 = 1$			
Stage 3: Screening record abstract			
Q_{31}	Does the abstract describe an analysis using administrative health database?	0	1
Q_{32}	Does the abstract describe an analysis conducted/related to humans?	0	1
Q_{33}	Does the abstract describe an analysis related to farming?	0	1
Q_{34}	Does the abstract describe an analysis related public health or epidemiological research?	0	1
$S_3 = Q_{31} \times Q_{32} \times Q_{33} \times Q_{34}$; Record eligible for stage 4 if $S_3 = 1$			
Stage 4: Screening record content		no	yes
Q_{41}	Does the record describe an analysis using administrative health database?	0	1
Q_{42}	Does the record describe an analysis conducted/related to humans?	0	1
Q_{43}	Does the record describe an analysis conducted/related to farming population?	0	1
Q_{44}	Does the record describe an analysis related public health or epidemiological research?	0	1
$S_4 = Q_{41} \times Q_{42} \times Q_{43} \times Q_{44}$; Record eligible for stage 4 if $S_4 = 1$			
Score = $S_1 \times S_2 \times S_3 \times S_4$; Record included and analyzed if score ≥ 1			

Note: Q: question, S: score.

126
 127

128 **Table S3.** Bibliometric index metrics.

129

130 • **h index**

131

132 The h index attempts to measure both the productivity and citation impact of the published body of work of an
133 entity (e.g., author, institution, journal) [1,2]. It refers to the total number of publications by a particular entity with
134 at least the same number of citations.

135

136 • **m index**

137

138 The m index is calculated by dividing the h index by the number of years of an entity's productive life (e.g.,
139 researcher) [1].

140

141 • **g index**

142

143 The g index of an entity corresponds to the largest number g such that the top g publications have at least g² or
144 more citations together [2].

145

146 • **Y index**

147

148 The Y index refers to the sum of both the total number of first-authored publications and the total number of
149 corresponding-author publications [3].

150

151 • **Dominance factor**

152

153 The dominance factor (DF) refers, for a particular researcher, to the proportion of multi-authored publications as
154 specific author's rank to the total number of multi-authored publications [4].

155

156 The dominance factor for being a first/primary author (DF_{first}) refers to the proportion of multi-authored
157 publications as a first author to the total number of multi-authored publications. It is calculated as follows:

158

159
$$DF_{first} = \frac{\text{number of multi – authored papers as first author}}{\text{total number of multi – authored papers}} \times 100$$

160

161 The dominance factor for being a last/senior author (DF_{last}) refers to the proportion of multi-authored publications
162 as a last author to the total number of multi-authored publications. It is calculated as follows:

163

164
$$DF_{last} = \frac{\text{number of multi – authored papers as last author}}{\text{total number of multi – authored papers}} \times 100$$

165

166 • **Fractionalized frequency**

167

168 The fractional frequency (FF) intends to reflect/measure an author's contribution. For one publication, it is
169 calculated as follows:

170

171
$$FF_i = \frac{1}{\text{total number of authors in the paper}_i}$$

172

173 For a given researcher, the FF corresponds to:

174

175 $FF_{author} = \sum_{i=1}^n FF_i$, where n refers to the total number of publications for which a researcher was an author.

176

177 • **Annual growth rate**

178

179 The annual growth rate refers to the variable's change in percentage as a year-over-year statistic [5], and is
180 calculated as follows:

181

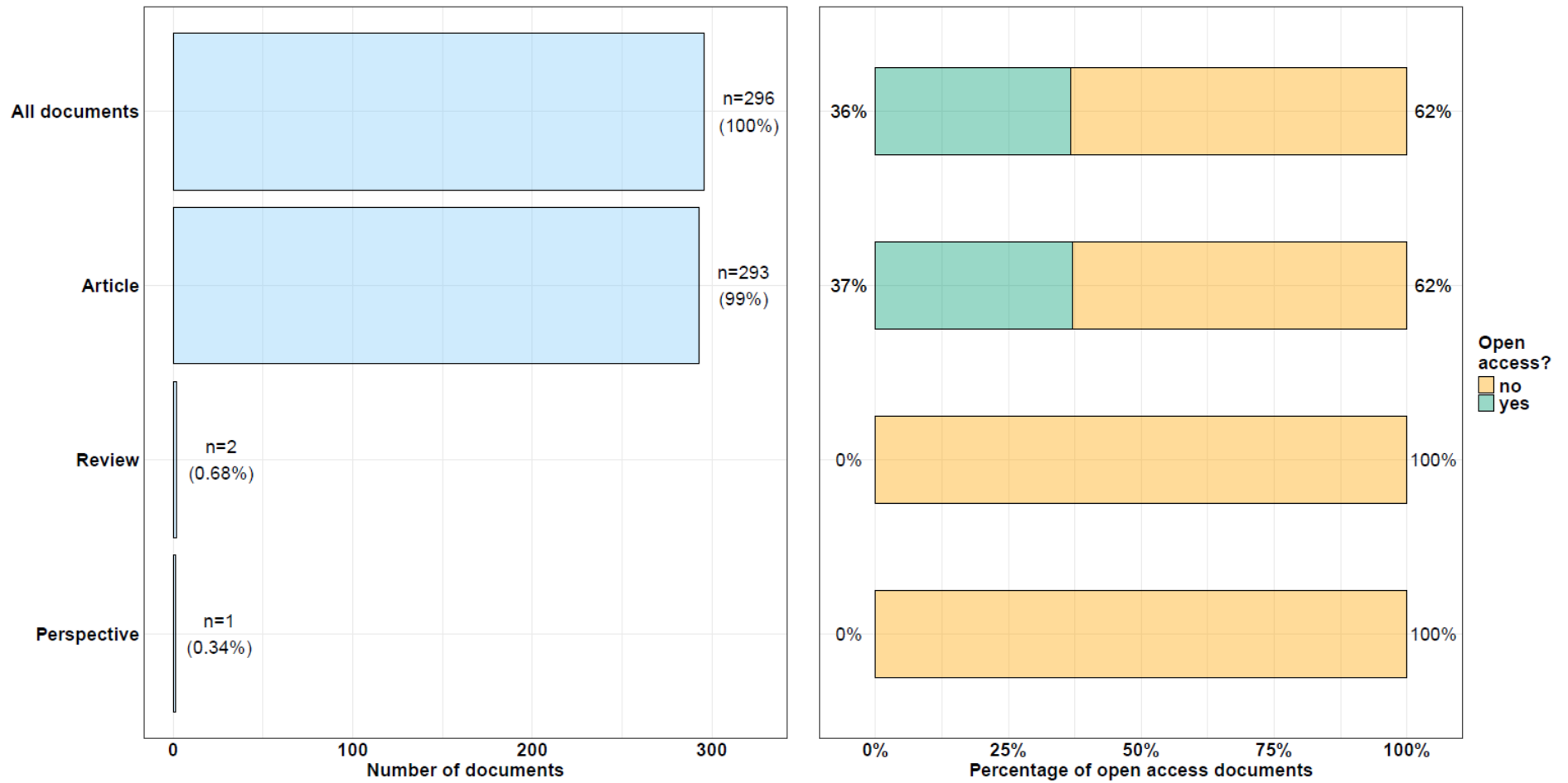
$$\text{Annual Growth Rate} = \frac{(\text{end value} - \text{first value})}{\text{first value}} \times 100$$

182 **Table S4.** The BIBLIO checklist for reporting the bibliometric reviews of the biomedical literature.
183

Section/Topic	Item No.	Checklist item	Reported on page No.
Title			
Identification	1	Identify the report as a bibliometric review in the title.	1
Issues/topics	2	Indicate the key issues/topics under investigation and coverage of time period.	1
Abstract			
Structured summary	3	Structured summary including (as applicable): background, methods, results (key findings) and conclusions.	2, 3
Introduction/ Background			
Justification/ Rationale/ Explanation	4	Present review of existing knowledge and epidemiological information.	4-7
Objectives	5	Statement of the objective (s) or question (s).	7
Methods			
Search engines (data sources)	6	Describe all information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources).	8-10
Search strategy	7	Keywords and systematization criteria (date of search, language, type of document) for the search.	8-11
Time period	8	The period that the review covers and the justification.	9
Eligibility criteria	9	Describe all inclusion and exclusion criteria; languages; study design, type of publication and time period.	9, 10
Data refinement (data selection procedure)	10	Remove the irrelevant articles; inspection to eliminate duplicate and unrelated articles (after evaluation of the title, abstract and content).	9, 10
Quality assessment (optional)	11	Assessment of papers by three authors and the use of assessing checklists.	NA
Data synthesis	12	Describe the methods used for summarizing, handling, synthesis, tabulations or schematic displays. Describe how the data were analyzed.	10-14
Results			
Descriptive findings (statistics)	13	- Provide details of the search and selection process in a flow diagram. - Number of citations retrieved (number of publication, year of publication, type of documents, country of publication, articles with the highest impact, most impactful authors, most impactful articles, authors with the highest production, top journals, top institutions, ...)	14-17, supplementary materials
Schematic map and trend	14	Summarize and/or present the schematic maps and trends using an appropriate software to present citations, journals, authors, top journals, time trends, emerging literature, and any relevant indicators (as applicable) [1-5].	27-39, supplementary materials
Tabulation and summarizing the findings	15	General recommendation: Studies under consideration could be summarized and organized by different subtitles and different scenarios. Regardless, results need to be presented in separate tables covering each subtitle. The followings are some options that could help to summarize the findings. <i>Option 1:</i> - Start the presentation with a historical view [when and who first published on the topic]. - Report on review papers. The result should be listed in a separate table. Also, specify the review type (scoping review, narrative review, systematic review, and meta-analysis). - Summarize the findings according to the study designs and main study types. <i>Option 2:</i> - Start the presentation with a historical view [when and who first published on the topic]. - Report on review papers. The result should be listed in a separate table. Also, indicate the review type (scoping review, narrative review, systematic review, and meta-analysis) should be specified. - Summarize the findings according to outcome measures or populations. For example, see [6]. <i>Option 3:</i> - Start the presentation with a historical view [when and who first published on the topic]. - Report on review papers. The result should be listed in a separate table. Also, specify the review type (scoping review, narrative review, systematic review, and meta-analysis). - Summarize the findings according to concept [7]. <i>Option 4:</i> - Start the presentation with a historical view [when and who first published on the topic]. - Report on review papers. The result should be listed in a separate table, and also specify the review type (scoping review, narrative review, systematic review, and meta-analysis). - Summarize the findings according to different subtitles relevant to the main topic [8].	14-39, supplementary materials
Synthesis of findings	16	Synthesize the findings as much as possible, find the gap, and propose a model, hypothesis, etc. (if applicable).	27-39, supplementary materials
Discussion			
Summary of evidence	17	Summarize the main findings. The findings should be presented in more "general" or "accessible" terms.	40-48
Interpretation	18	Include interpretation consistent with results. Explanations for observed outcomes, similarities, and differences reported would be essential.	40-48
Strengths and limitations	19	Discuss the strengths and limitations.	48-49
Conclusion(s)	20	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications.	49-51
<p>1. McDougal L, Dehingia N, Cheung WW, Dixit A, Raj A. COVID-19 burden, author affiliation and women's well-being: A bibliometric analysis of COVID-19 related publications including focus on low-and middle-income countries. <i>eClinicalMedicine</i> 2022; 52: 101606.</p> <p>2. Henstock L, Wong R, Tsuchiya A, Spencer A. Behavioral theories that have influenced the way health state preferences are elicited and interpreted: A bibliometric mapping analysis of the time trade-off method with VOSviewer visualization. <i>Front Health Serv</i> 2022; 2: 848087.</p> <p>3. Bodea F, Bungau SG, Negru AP, Radu A, Tarce AG, Tit DM, et al. Exploring new therapeutic avenues for ophthalmic disorders: Glaucoma-related molecular docking evaluation and bibliometric analysis for improved management of ocular diseases. <i>Bioengineering</i> 2023; 10(8): 983.</p> <p>4. Sang XZ, Wang CQ, Chen W, Rong H, Hou LJ. An exhaustive analysis of post-traumatic brain injury dementia using bibliometric methodologies. <i>Front Neurol</i> 2023; 14: 1165059.</p> <p>5. Ramli MI, Hamzaid NA, Engkasan JP, Usman J. Respiratory muscle training: a bibliometric analysis of 60 years' multidisciplinary journey. <i>Biomed Eng Online</i> 2023; 22(1): 50.</p> <p>6. Akosman I, Kumar N, Mortenson R, Lans A, De La Garza Ramos R, Eleswarapu A, et al. Racial differences in perioperative complications, readmissions, and mortalities after elective spine surgery in the United States: A systematic review using AI-assisted bibliometric analysis. <i>Glob Spine J</i> 2023; 21925682231186759.</p> <p>7. Tavousi M, Mohammadi S, Sadighi J, Zarei F, Kermani RM, Rostami R, Montazeri A. Measuring health literacy: A systematic review and bibliometric analysis of instruments from 1993 to 2021. <i>Plos One</i> 2022; 17(7): e0271524.</p> <p>8. Montazeri A. Health-related quality of life in breast cancer patients: A bibliographic review of the literature from 1974 to 2007. <i>J Exp Clin Cancer Res</i> 2008; 27: 32.</p>			

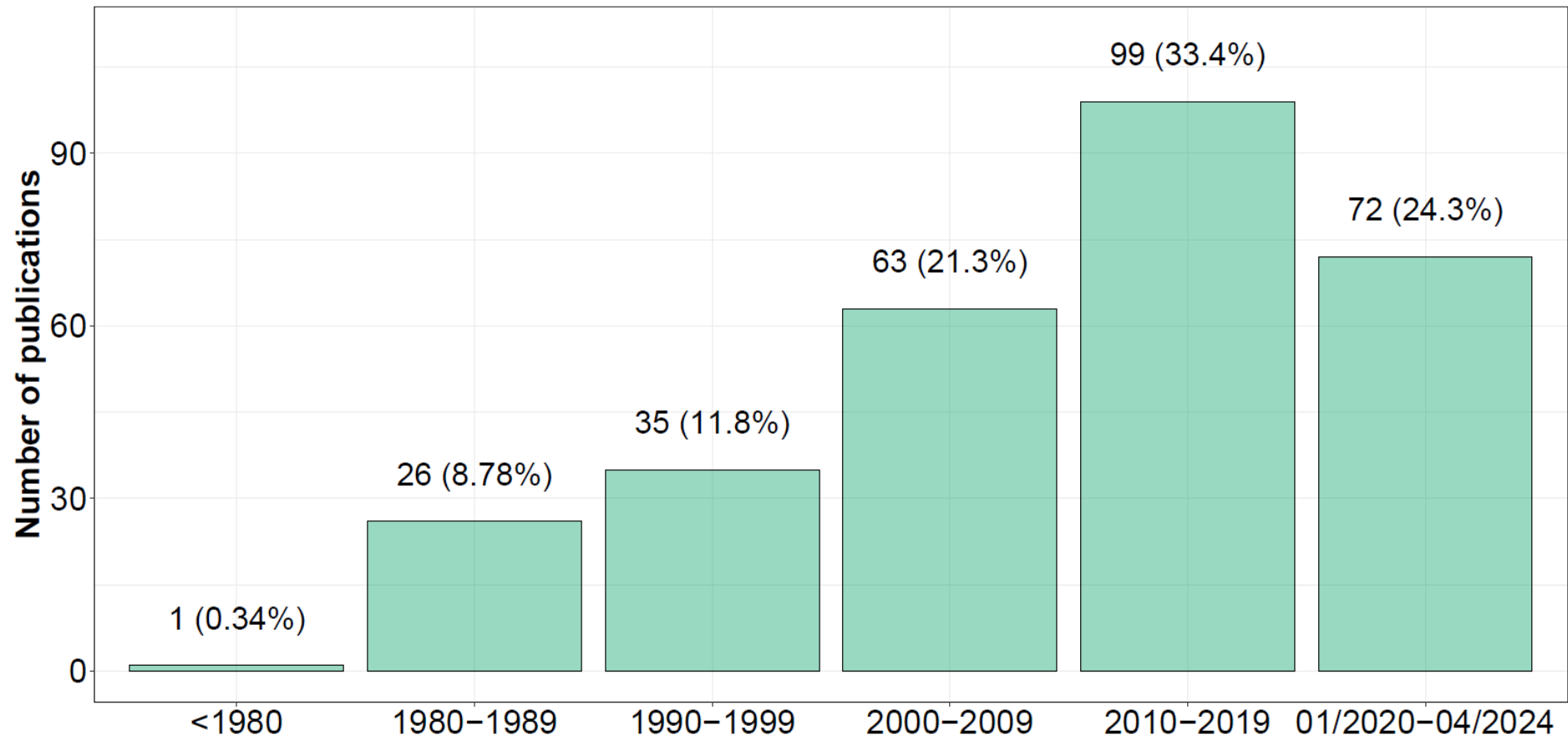
184 **Rights and permissions:** The original source of the checklist is: Montazeri A, Mohammadi S, M.Hesari P, Ghaemi M, Riazi H,
185 Sheikhi-Mobarakeh Z. Preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO): a minimum
186 requirements. *Systematic Reviews* 2023; 12: 239. doi: 10.1186/s13643-023-02410-2. The article is licensed under a Creative Commons
187 Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). A changes was made to the original checklist to add in full
188 references to the cited sources.

189 **Figure S1.** Publication type.



190

191 **Figure S2.** Scientific productivity in terms of publications per period of time.



192

193 **Figure S3.** Historical direct citation network.



194

195

196 Jouneau S, Marette S, Robert AM, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in dairy farmers: AIRBAg study. Environ Res. 2019;169:1-6.

197 doi:10.1016/j.envres.2018.10.026

198 Roux P, Guillien A, Soumagne T, et al. Smoking habits in French farmers: a cross-sectional study. BMC Public Health. 2017;17(1):166. doi:10.1186/s12889-017-4030-4

199 Alavanja MC, Dosemeci M, Samanic C, et al. Pesticides and lung cancer risk in the agricultural health study cohort. Am J Epidemiol. 2004;160(9):876-885.

200 doi:10.1093/aje/kwh290

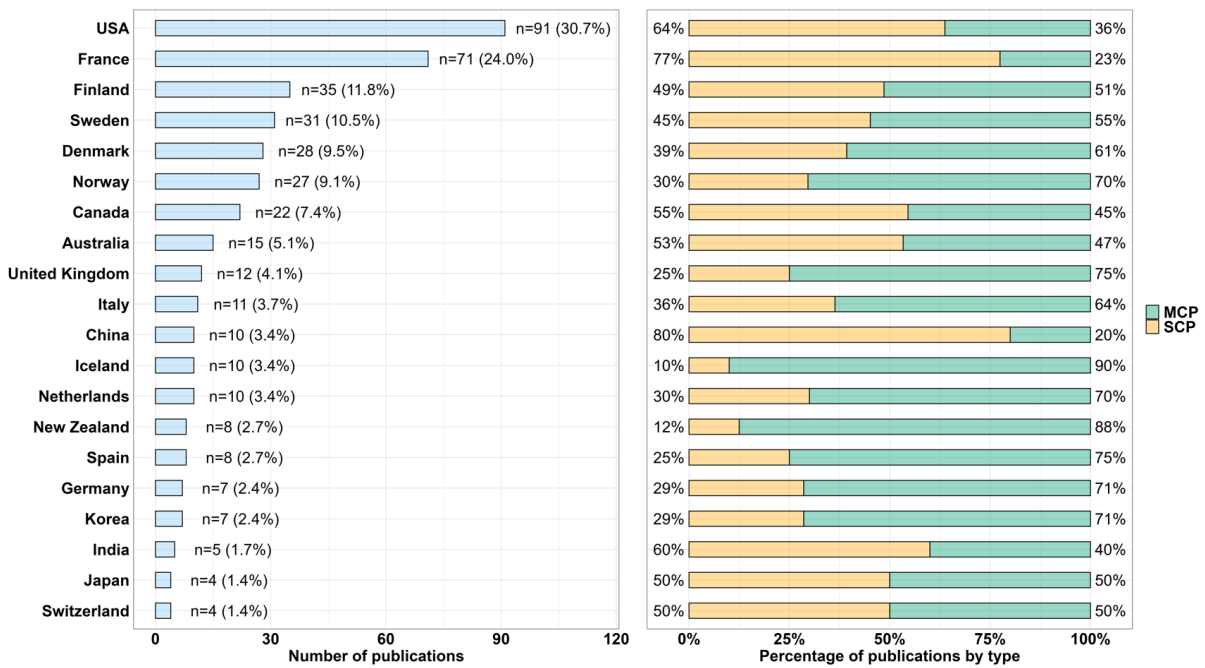
201 Reif J, Pearce N, Fraser J. Cancer risks in New Zealand farmers. *Int J Epidemiol.* 1989;18(4):768-774. doi:10.1093/ije/18.4.768
202 Marescaux A, Degano B, Soumagne T, Thaon I, Laplante JJ, Dalphin JC. Impact of farm modernity on the prevalence of chronic obstructive pulmonary disease in dairy farmers.
203 *Occup Environ Med.* 2016;73(2):127-133. doi:10.1136/oemed-2014-102697
204 Tual S, Lemarchand C, Boulanger M, et al. Exposure to Farm Animals and Risk of Lung Cancer in the AGRICAN Cohort. *Am J Epidemiol.* 2017;186(4):463-472.
205 doi:10.1093/aje/kwx125.
206 Petit P, Gandon G, Chabardès S, Bonnetterre V. Agricultural activities and risk of central nervous system tumors among French farm managers: Results from the TRACTOR
207 project. *Int J Cancer.* 2022;151(10):1737-1749. doi:10.1002/ijc.34197.
208 Piel C, Pouchieu C, Tual S, et al. Central nervous system tumors and agricultural exposures in the prospective cohort AGRICAN. *Int J Cancer.* 2017;141(9):1771-1782.
209 doi:10.1002/ijc.30879
210 Lemarchand C, Tual S, Levêque-Morlais N, et al. Cancer incidence in the AGRICAN cohort study (2005-2011). *Cancer Epidemiol.* 2017;49:175-185.
211 doi:10.1016/j.canep.2017.06.003.
212 Morrison H, Savitz D, Semenciw R, et al. Farming and prostate cancer mortality. *Am J Epidemiol.* 1993;137(3):270-280. doi:10.1093/oxfordjournals.aje.a116674.
213 Alavanja MC, Samanic C, Dosemeci M, et al. Use of agricultural pesticides and prostate cancer risk in the Agricultural Health Study cohort. *Am J Epidemiol.* 2003;157(9):800-
214 814. doi:10.1093/aje/kwg040
215 Mills PK, Yang R. Prostate cancer risk in California farm workers. *J Occup Environ Med.* 2003;45(3):249-258. doi:10.1097/01.jom.0000058339.05741.0c.
216 Kachuri L, Harris MA, MacLeod JS, Tjepkema M, Peters PA, Demers PA. Cancer risks in a population-based study of 70,570 agricultural workers: results from the Canadian
217 census health and Environment cohort (CanCHEC). *BMC Cancer.* 2017;17(1):343. doi:10.1186/s12885-017-3346-x
218 Togawa K, Leon ME, Lebailly P, et al. Cancer incidence in agricultural workers: Findings from an international consortium of agricultural cohort studies (AGRICOH). *Environ*
219 *Int.* 2021;157:106825. doi:10.1016/j.envint.2021.106825.
220 Pukkala E, Martinsen JI, Lynge E, et al. Occupation and cancer - follow-up of 15 million people in five Nordic countries. *Acta Oncol.* 2009;48(5):646-790.
221 doi:10.1080/02841860902913546
222 Leon ME, Schinasi LH, Lebailly P, et al. Pesticide use and risk of non-Hodgkin lymphoid malignancies in agricultural cohorts from France, Norway and the USA: a pooled
223 analysis from the AGRICOH consortium. *Int J Epidemiol.* 2019;48(5):1519-1535. doi:10.1093/ije/dyz017
224 Bucchi L, Nanni O, Ravaioli A, et al. Cancer mortality in a cohort of male agricultural workers from northern Italy. *J Occup Environ Med.* 2004;46(3):249-256.
225 doi:10.1097/01.jom.0000116804.10496.fc
226 Brownson RC, Chang JC, Davis JR, Bagby JR Jr. Occupational risk of prostate cancer: a cancer registry-based study. *J Occup Med.* 1988;30(6):523-526.

227 **Figure S4.** Most prolific authors' country.

228

229 MPC: multi-country publication, SCP: single-country publication.

230



231

232 **Table S5.** Top 25 of the most cited publications.

Publication	Publication year	Journal	DOI	Total citation	Local citation (%)	Average citation per year
Pukkala et al. 2009 [6]	2009	Acta Oncol	10.1080/02841860902913546	485	21 (4.33%)	30.3
Alavanja et al. 2003 [7]	2003	Am J Epidemiol	10.1093/aje/kwg040	283	6 (2.12%)	12.9
Elbaz et al. 2009 [8]	2009	Ann Neurol	10.1002/ana.21717	206	7 (3.40%)	12.9
Andersen et al. 1999 [9]	1999	Scand J Work Environ Health	none	206	0 (0.00%)	7.92
Karjalainen et al. 2001 [10]	2001	Am J Respir Crit Care Med	10.1164/ajrccm.164.4.2012146	162	0 (0.00%)	6.75
Vingård et al. 1991 [11]	1991	Int J Epidemiol	10.1093/ije/20.4.1025	157	1 (0.64%)	4.62
Alavanja et al. 2004 [12]	2004	Am J Epidemiol	10.1093/aje/kwh290	155	7 (4.52%)	7.38
Andreotti et al. 2018 [13]	2018	J Natl Cancer Inst	10.1093/jnci/djx233	143	0 (0.00%)	20.4
Wigle et al. 1990 [14]	1990	J Natl Cancer Inst	10.1093/jnci/82.7.575	136	0 (0.00%)	3.89
Engel et al. 2005 [15]	2005	Am J Epidemiol	10.1093/aje/kwi022	116	2 (1.72%)	5.8
Hemminki et al. 1981 [16]	1981	J Epidemiol Community Health	10.1136/jech.35.1.11	114	0 (0.00%)	2.59
Weichenthal et al. 2014 [17]	2014	Environ Health Perspect	10.1289/ehp.1307277	113	0 (0.00%)	10.3
Levecque et al. 2003 [18]	2003	Hum Mol Genet	10.1093/hmg/ddg009	111	0 (0.00%)	5.05
Morrison et al. 1993 [19]	1993	Am J Epidemiol	10.1093/oxfordjournals.aje.a116674	106	5 (4.72%)	3.31
Kristensen et al. 1996 [20]	1996	Scand J Work Environ Health	10.5271/sjweh.104	98	0 (0.00%)	3.38
Orton et al. 2011 [21]	2011	Neurology	10.1212/WNL.0b013e31820a0a9f	94	1 (1.06%)	6.71
Stokes et al. 1995 [22]	1995	Occup Environ Med	10.1136/oem.52.10.648	92	0 (0.00%)	3.07
Vukusic et al. 2007 [23]	2007	J Neurol Neurosurg Psychiatry	10.1136/jnnp.2006.101196	90	0 (0.00%)	5.00
Mills and Yang, 2003 [24]	2003	J Occup Environ Med	10.1097/01.jom.0000058339.05741.0c	88	5 (5.68%)	4.00
Eriksson and Karlsson, 1992 [25]	1992	Br J Ind Med	none	87	0 (0.00%)	2.64
Wiklund et al. 1989 [26]	1989	Br J Ind Med	none	87	0 (0.00%)	2.42
Pouchieu et al. 2018 [27]	2018	Int J Epidemiol	10.1093/ije/dyx225	86	1 (1.16%)	12.3
Karjalainen et al. 2003 [28]	2003	Chest	10.1378/chest.123.1.283	85	0 (0.00%)	3.86
Leon et al. 2019 [29]	2019	Int J Epidemiol	10.1093/ije/dyz017	85	5 (5.88%)	14.2
Lee et al. 2007 [30]	2007	Int J Cancer	10.1002/ijc.22635	83	2 (2.41%)	4.61
Reif et al. 1989 [31]	1989	Int J Epidemiol	10.1093/ije/18.4.768	83	4 (4.82%)	2.31

233 *Note:* DOI: digital object identifier. Please refer to the beginning of the supplemental materials for the definition of the bibliometric indices.

234 **Table S6.** Other bibliometric analysis information.

235

- 236 • *The most prolific countries*

237

238 Studies were led by authors from 34 countries, predominantly high-income nations, with 24.3% of studies (72/296)
 239 involving multi-country collaborations. US-based authors contributed the most publications (91/296, 30.7%),
 240 followed by France (71/296, 24.0%) and Finland (35/296, 11.8%). US authors also had the most citations
 241 (3495/9379, 37.2%), with France and Finland ranking second and third. The highest average AGRs were seen in
 242 authors from India (25%), France (15%), and the US (10.1%) (Table MA1).

243

244 Positive associations were found between the number of publications and GDP ($\rho=0.46$, $p=5.9e-3$), R&D
 245 researchers ($\rho=0.62$, $p=1.1e-4$), and cereal yield ($\rho=0.41$, $p=0.02$). In contrast, negative associations were observed
 246 with agricultural employment ($\rho=-0.60$, $p=1.6e-4$), sex-specific employment in agriculture, agricultural value
 247 added to GDP ($\rho=-0.51$, $p=2.1e-3$), as well as fertilizer consumption ($\rho=-0.37$, $p=0.04$) (Table MA2).

248

249 Countries with the highest scientific production per GDP were Iceland, Finland, and Malawi, while Iceland,
 250 Finland, and Norway led per capita and cereal production (Table MA3). The US, Norway, and Canada exhibited
 251 the highest scientific production per fertilizer consumption, and Norway, Finland, and Denmark excelled in
 252 scientific output per hectare of agricultural land. The US, France, and Finland were leaders in scientific output
 253 related to crop, food, livestock production indices, and cereal yield. Notably, the US, Sweden, and France led in
 254 output per agricultural employment, while the US, Canada, and the UK had the highest scientific production
 255 relative to agriculture, forestry, and fishing value added to GDP.

256

257 **Table MA1.** The most productive author's countries.

Country	Continent	Publication period	Number of publications (%)	Total citation	Local citation (%)	Average citation per year	Mean AGR	Fractionalized frequency (%)	h-index	g-index	m-index
USA	North America	1981-2024	91 (30.7%)	3495	81 (2.32%)	81.3	10.1	72.6	37	57	0.841
France	Europe	1998-2024	71 (24.0%)	1762	69 (3.92%)	65.3	15.0	60.3	22	40	0.815
Finland	Europe	1980-2024	35 (11.8%)	1429	30 (2.10%)	37.2	-16.7	21.6	17	35	0.378
Sweden	Europe	1986-2024	31 (10.5%)	1290	33 (2.56%)	43.9	-17.5	18.2	19	31	0.487
Denmark	Europe	1988-2024	28 (9.46%)	828	31 (3.74%)	31.9	-9.46	15.1	15	28	0.405
Norway	Europe	1996-2024	27 (9.12%)	965	37 (3.83%)	41.8	-9.20	12.0	13	27	0.448
Canada	North America	1975-2022	22 (7.43%)	799	13 (1.63%)	18.0	-13.9	16.3	13	22	0.260
Australia	Oceania	1986-2024	15 (5.07%)	177	9 (5.08%)	4.85	-6.41	10.6	7	13	0.179
UK	Europe	1984-2021	12 (4.05%)	294	9 (3.06%)	8.50	-15.8	6.53	9	12	0.220
Italy	Europe	1988-2024	11 (3.72%)	170	7 (4.12%)	4.59	-9.01	6.58	7	11	0.189
China	Asia	1995-2024	10 (3.38%)	85	0 (0.00%)	3.73	-16.4	9.00	4	10	0.114
Iceland	Europe	2006-2024	10 (3.38%)	570	24 (4.21%)	30.0	-21.1	2.74	7	10	0.368
Netherlands	Europe	2004-2023	10 (3.38%)	217	14 (6.45%)	10.9	-18.3	4.64	6	10	0.286
New Zealand	Oceania	1986-2021	8 (2.70%)	364	11 (3.02%)	10.1	-5.56	3.86	7	8	0.179
Spain	Europe	1998-2023	8 (2.70%)	186	4 (2.15%)	7.31	-23.1	4.32	6	8	0.222
Germany	Europe	2001-2024	7 (2.36%)	166	2 (1.20%)	6.92	-16.7	3.79	4	7	0.167
South Korea	Asia	2002-2021	7 (2.36%)	136	7 (5.15%)	6.80	-20.0	4.11	5	7	0.217
India	Asia	2020-2023	5 (1.69%)	22	1 (4.55%)	5.50	25.0	4.00	2	4	0.400
Japan	Asia	1999-2023	4 (1.35%)	59	0 (0.00%)	2.36	-8.00	3.00	3	4	0.115
Switzerland	Europe	1988-2014	4 (1.35%)	98	5 (5.10%)	7.26	-11.1	2.62	4	4	0.108
Belgium	Europe	2016-2017	2 (0.68%)	15	0 (0.00%)	7.50	0.00	1.33	2	2	0.222
Brazil	South America	2017-2024	2 (0.68%)	5	0 (0.00%)	0.63	-12.5	1.50	1	2	0.125
Ethiopia	Africa	2022-2024	2 (0.68%)	1	0 (0.00%)	0.33	-33.3	1.50	1	1	0.333
Austria	Europe	2019	1 (0.34%)	1	0 (0.00%)	1.00	0.00	0.33	1	1	0.167
Colombia	South America	2020	1 (0.34%)	15	0 (0.00%)	15.0	0.00	0.25	1	1	0.200
Costa Rica	Central America	2006	1 (0.34%)	12	0 (0.00%)	12.0	0.00	0.50	1	1	0.053
Greece	Europe	2016	1 (0.34%)	7	0 (0.00%)	7.00	0.00	0.33	1	1	0.111
Laos	Asia	2019	1 (0.34%)	6	0 (0.00%)	6.00	0.00	0.50	1	1	0.167
Malawi	Africa	2015	1 (0.34%)	15	0 (0.00%)	15.0	0.00	0.50	1	1	0.100
Malaysia	Asia	2001	1 (0.34%)	24	0 (0.00%)	24.0	0.00	0.50	1	1	0.042
Mexico	Central America	2020	1 (0.34%)	15	0 (0.00%)	15.0	0.00	0.25	1	1	0.200
Morocco	Africa	2018	1 (0.34%)	7	0 (0.00%)	7.00	0.00	0.50	1	1	0.143
Nigeria	Africa	2018	1 (0.34%)	1	0 (0.00%)	1.00	0.00	1.00	1	1	0.143
Turkey	Europe	2021	1 (0.34%)	0	0 (0.00%)	0.00	0.00	0.50	1	1	0.250

258 *Note:* AGR: annual growth rate. Please refer to the beginning of the supplemental materials for the definition of
 259 the bibliometric indices.

260

Table MA2. Spearman correlation between scientific production and country characteristics

Country characteristics	Spearman 's ρ	p
Scientific production per researchers in R&D (per million people)	0.62	1.1e-04
Scientific production per GDP	0.46	5.9e-03
Scientific production per cereal yield (kg per hectare)	0.41	0.016
Scientific production per female employment in agriculture (% of female employment)	-0.69	6.6e-06
Scientific production per employment in agriculture (% of total employment)	-0.66	2.1e-05
Scientific production per male employment in agriculture (% of male employment)	-0.60	1.6e-04
Scientific production per permanent cropland (% of land area)	-0.53	1.5e-03
Scientific production per agriculture, forestry, and fishing, value added (% of GDP)	-0.51	2.1e-03
Scientific production per fertilizer consumption (% of fertilizer production)	-0.37	0.036
Scientific production per fertilizer consumption (kilograms per hectare of arable land)	-0.37	0.036
Scientific production per livestock production index	-0.26	0.14
Scientific production per food production index	-0.24	0.17
Scientific production per cereal production (metric tons)	0.20	0.25
Scientific production per crop production index	-0.14	0.44
Scientific production per land under cereal production (hectares)	0.12	0.48
Scientific production per population	-0.07	0.69
Scientific production per agricultural land (% of land area)	-0.05	0.77
Scientific production per total labor force	-0.03	0.85
Scientific production per agricultural land (km ²)	0.01	0.96

261

262

Note: GDP: gross domestic product, R&D: research and development.

Table MA3. Spearman correlation between scientific production and country characteristics

Country	SciProd per GDP	SciProd per pop.	SciProd per res. in R&D (per million people)	SciProd per fertilizer cons. (% of fertilizer prod.)	SciProd per fertilizer cons. (kg per hectare of arable land)	SciProd per agricult. land (km ²)	SciProd per agricult. land (% of Land area)	SciProd per land under cereal prod. (hectares)	SciProd per permanent cropland (% of land area)	SciProd per cereal prod. (metric tons)	SciProd per crop prod. index	SciProd per food prod. index	SciProd per livestock prod. index	SciProd per cereal yield (kg per hectare)	SciProd per female employ. in agri. (% of female employ.)	SciProd per male employ. in agri. (% of male employ.)	SciProd per employ. in agri. (% of total employ.)	SciProd per agri., forestry, and fishing, value added (% of GDP)	SciProd per total labor force
USA	3.6e-12	2.7e-07	0.02	0.67	0.67	2.0e-05	1.86	1.2e-06	3.1e+02	1.8e-07	0.85	0.87	0.82	0.01	59.3	22.8	32.8	68.1	5.3e-07
France	2.6e-11	1.1e-06	0.01	0.05	0.21	2.1e-04	1.13	7.2e-06	21.6	9.7e-07	0.67	0.66	0.63	9.4e-03	15.1	10.9	12.4	6.74	2.2e-06
Finland	1.3e-10	6.3e-06	4.5e-03	0.17	0.15	1.2e-03	3.69	2.6e-05	2.1e+03	8.2e-06	0.30	0.31	0.31	8.2e-03	5.85	3.32	4.19	3.57	1.2e-05
Sweden	5.3e-11	3.0e-06	3.9e-03	0.10	0.17	7.3e-04	2.98	1.9e-05	2.8e+03	4.5e-06	0.26	0.27	0.27	4.9e-03	16.6	5.99	8.83	6.17	5.3e-06
Denmark	7.1e-11	4.7e-06	3.6e-03	0.06	0.10	8.9e-04	0.35	1.5e-05	36.1	2.8e-06	0.26	0.27	0.26	3.9e-03	9.98	3.64	5.15	4.19	8.8e-06
Norway	4.7e-11	5.0e-06	4.0e-03	0.52	0.08	2.4e-03	8.71	7.3e-05	2.0e+03	1.7e-05	0.21	0.26	0.26	5.6e-03	7.54	3.54	4.67	5.29	9.0e-06
Canada	1.0e-11	5.7e-07	4.6e-03	0.30	0.17	3.5e-05	3.10	9.9e-07	1.1e+03	3.3e-07	0.20	0.20	0.19	5.2e-03	9.21	4.99	6.30	10.0	1.0e-06
Australia	9.0e-12	5.8e-07	3.3e-03	0.04	0.14	2.9e-06	0.22	7.3e-07	2.9e+02	2.7e-07	0.10	0.13	0.15	5.2e-03	3.91	2.33	2.80	3.58	1.1e-06
UK	3.9e-12	1.8e-07	2.6e-03	0.03	0.03	6.1e-05	0.15	3.0e-06	23.4	4.5e-07	0.11	0.12	0.11	1.5e-03	9.75	4.07	5.54	9.41	3.5e-07
Italy	5.5e-12	1.9e-07	4.1e-03	0.01	0.04	5.3e-05	0.16	1.7e-06	1.03	4.7e-07	0.08	0.09	0.10	1.9e-03	1.25	1.33	1.30	3.40	4.2e-07
China	5.6e-13	7.1e-09	6.3e-03	0.05	0.02	1.9e-06	0.18	9.7e-08	4.94	1.6e-08	0.09	0.09	0.09	1.6e-03	0.19	0.16	0.17	0.24	1.3e-08
Iceland	3.6e-10	2.6e-05	1.3e-03	4.8e-04	0.04	4.7e-04	0.47	3.3e-03	no data	1.1e-03	0.08	0.09	0.09	2.1e-03	2.02	0.68	0.98	1.14	4.3e-05
Netherlands	1.0e-11	5.7e-07	1.7e-03	0.12	0.01	4.4e-04	0.15	1.9e-05	6.89	4.6e-06	0.09	0.10	0.10	1.1e-03	3.54	1.92	2.33	1.78	1.0e-06
New Zealand	3.2e-11	1.6e-06	1.4e-03	0.02	3.6e-03	5.3e-05	0.14	3.0e-05	28.5	6.7e-06	0.07	0.08	0.08	8.9e-04	1.04	0.60	0.74	0.61	2.6e-06
Spain	5.7e-12	1.7e-07	2.6e-03	0.06	0.05	2.4e-05	0.12	1.0e-06	0.79	2.9e-07	0.07	0.07	0.07	1.8e-03	0.90	0.70	0.75	1.82	3.4e-07
Germany	1.7e-12	8.3e-08	1.3e-03	0.10	0.02	3.6e-05	0.13	8.9e-07	4.83	1.4e-07	0.07	0.07	0.06	8.7e-04	1.93	1.87	1.89	6.35	1.6e-07
South Korea	3.9e-12	1.4e-07	8.0e-04	7.7e-03	0.01	3.0e-04	0.29	3.0e-06	3.05	6.6e-07	0.06	0.07	0.06	1.0e-03	0.38	0.47	0.43	0.15	2.4e-07
India	1.5e-12	3.5e-09	0.02	0.02	0.02	2.8e-06	0.08	4.7e-08	1.09	1.4e-08	0.04	0.04	0.04	1.4e-03	0.06	0.09	0.08	0.12	8.4e-09
Japan	9.5e-13	3.2e-08	7.3e-04	0.03	8.3e-03	5.6e-05	0.21	8.2e-07	2.30	1.9e-07	0.03	0.04	0.04	5.9e-04	0.51	0.67	0.59	2.11	5.8e-08
Switzerland	5.0e-12	4.6e-07	7.2e-04	3.9e-03	8.2e-03	2.3e-04	0.09	1.9e-05	6.19	2.8e-06	0.03	0.04	0.04	5.7e-04	0.98	0.66	0.79	1.90	8.0e-07
Belgium	3.5e-12	1.7e-07	3.5e-04	0.03	3.4e-03	1.1e-04	0.04	5.7e-06	2.52	6.1e-07	0.02	0.02	0.02	2.0e-04	0.92	0.59	0.69	1.49	3.7e-07
Brazil	1.0e-12	9.3e-09	2.3e-03	2.5e-03	5.4e-03	8.3e-07	0.07	7.2e-08	1.43	1.5e-08	0.02	0.02	0.02	3.8e-04	0.18	0.08	0.11	0.12	1.8e-08
Ethiopia	1.6e-11	1.6e-08	0.02	no data	0.05	3.4e-06	0.04	1.8e-07	0.99	6.3e-08	0.02	0.02	0.02	6.9e-04	0.03	0.02	0.03	0.03	3.2e-08
Austria	2.1e-12	1.1e-07	1.7e-04	7.0e-03	3.6e-03	2.8e-05	0.02	9.4e-07	0.83	1.7e-07	8.3e-03	9.7e-03	9.9e-03	1.4e-04	0.12	0.15	0.14	0.20	2.1e-07
Colombia	2.9e-12	1.9e-08	0.01	7.0e-04	1.4e-03	2.2e-06	0.02	5.7e-07	0.44	2.2e-07	9.0e-03	9.0e-03	9.1e-03	2.3e-04	0.13	0.03	0.05	0.04	3.9e-08
Costa Rica	1.5e-11	1.9e-07	1.8e-03	4.3e-05	1.1e-03	3.6e-05	0.02	6.1e-06	0.14	2.6e-06	9.5e-03	9.5e-03	8.8e-03	2.1e-04	0.10	0.04	0.05	0.04	3.9e-07
Greece	4.6e-12	9.5e-08	2.5e-04	1.6e-03	4.1e-03	1.1e-05	0.01	5.6e-07	0.11	1.6e-07	8.7e-03	9.2e-03	8.6e-03	2.2e-04	0.04	0.06	0.05	0.14	2.0e-07
Laos	6.5e-11	1.3e-07	0.06	no data	0.02	4.3e-05	0.10	8.1e-07	1.37	1.8e-07	8.8e-03	8.7e-03	5.5e-03	2.2e-04	0.01	0.01	0.01	0.02	3.2e-07
Malawi	7.6e-11	4.9e-08	0.02	1.0e-05	0.01	1.7e-05	0.02	4.7e-07	0.47	2.0e-07	7.4e-03	6.8e-03	5.3e-03	4.1e-04	0.01	0.01	0.01	0.03	1.2e-07
Malaysia	2.5e-12	3.0e-08	4.2e-04	5.0e-04	4.4e-04	1.2e-05	0.04	1.3e-06	0.04	3.4e-07	9.4e-03	9.7e-03	9.5e-03	2.2e-04	0.07	0.05	0.05	0.02	5.8e-08
Mexico	7.1e-13	7.8e-09	2.4e-03	1.9e-03	9.2e-03	9.4e-07	0.02	8.7e-08	0.69	2.6e-08	8.7e-03	8.5e-03	8.3e-03	2.5e-04	0.09	0.03	0.04	0.08	1.7e-08
Morocco	7.5e-12	2.7e-08	9.3e-04	3.5e-03	0.01	3.2e-06	0.01	1.6e-07	0.25	8.6e-08	8.4e-03	8.6e-03	8.6e-03	4.3e-04	0.02	0.02	0.02	0.04	8.1e-08
Nigeria	2.1e-12	4.6e-09	0.03	1.6e-04	0.05	1.5e-06	0.01	5.2e-08	0.14	3.3e-08	8.3e-03	8.3e-03	8.9e-03	5.8e-04	0.02	0.02	0.02	0.03	1.3e-08
Turkey	1.1e-12	1.2e-08	5.6e-04	2.1e-03	6.7e-03	2.4e-06	0.02	7.1e-08	0.21	2.6e-08	8.0e-03	7.8e-03	7.3e-03	2.9e-04	0.01	0.03	0.02	0.02	2.9e-08

264 Note: agri.: agriculture, agricult.: agricultural, cons.: consumption, employ.: employment, GDP: gross domestic product, pop.: population, prod.: production, R&D: research and
265 development, res.: researcher, SciProd: scientific production.

266 • *The most active journals*

267

268 Articles were published in 118 journals (Table 4). Most journals published less than five studies and received on average 3.86 citations. The most active journals were the
 269 *American Journal Of Industrial Medicine* (18/296, 6.1%), the *Journal Of Occupational And Environmental Medicine* (14/296, 4.7%), as well as the *International Journal Of*
 270 *Epidemiology*, the *Occupational And Environmental Medicine*, and the *Scandinavian Journal Of Work Environment & Health*, which all had 13 publications (Table MA4).
 271 Notably, 11 of the top 20 journals are ranked in the first quartile (Q1) of their respective category, primarily in public health and occupational health, with 2022 impact factors
 272 ranging from 2.3 to 12.7.

273

274 **Table MA4.** Top 20 of the most active journals.

Journal	Publication start	Number of publications (%)	Total citation	h-index	Average citation per year	Mean AGR	g-index	m-index	Impact factor (2022)	Rank
American Journal Of Industrial Medicine	1989	18 (6.08%)	515	15	14.3	-27.8	18	0.42	3.5	Q2
Journal Of Occupational And Environmental Medicine	1985	14 (4.73%)	423	12	10.6	-20.8	14	0.30	3.2	Q3
International Journal Of Epidemiology	1980	13 (4.39%)	813	13	18.1	-14.2	13	0.29	7.7	Q1
Occupational And Environmental Medicine	1995	13 (4.39%)	351	9	11.7	-20.0	13	0.30	4.9	Q1
Scandinavian Journal Of Work Environment & Health	1990	13 (4.39%)	649	12	18.5	-20.6	13	0.34	6.3	Q1
Cancer Causes & Control	1997	9 (3.04%)	223	8	7.96	-14.8	9	0.29	2.3	Q4
Environmental Research	1993	9 (3.04%)	76	4	2.38	-6.25	8	0.13	8.3	Q1
Environmental Health Perspectives	2006	7 (2.36%)	357	7	18.8	-26.7	7	0.37	10.5	Q1
International Journal Of Environmental Research And Public Health	2011	7 (2.36%)	56	3	4.00	-11.5	7	0.21	4.6	Q2
International Archives Of Occupational And Environmental Health	2001	6 (2.03%)	118	6	4.92	-23.5	6	0.25	3.0	Q3
International Journal Of Cancer	1989	6 (2.03%)	153	4	4.25	-11.8	6	0.11	6.4	Q1
Journal Of Agromedicine	2009	6 (2.03%)	98	5	6.13	12.5	6	0.31	2.4	Q3
American Journal Of Epidemiology	1993	5 (1.69%)	669	5	20.9	-8.00	5	0.16	5.0	Q1
Journal Of Epidemiology And Community Health	1981	5 (1.69%)	264	5	6.00	-31.2	5	0.11	6.3	Q1
British Journal Of Industrial Medicine	1988	5 (1.69%)	256	5	6.92	-30	5	0.14	none	none
BMJ Open	2013	4 (1.35%)	35	3	2.92	-28.6	4	0.25	2.9	Q2
Frontiers In Public Health	2022	4 (1.35%)	8	2	2.67	0.00	2	0.67	5.2	Q1
American Journal Of Public Health	1990	3 (1.01%)	144	3	4.11	-5.00	3	0.09	12.7	Q1
Annals Of Epidemiology	2010	3 (1.01%)	133	3	8.87	-22.2	3	0.20	5.6	Q1
Cancer Epidemiology	2017	3 (1.01%)	40	2	5.00	-14.3	3	0.25	2.6	Q2

275 *Note:* AGR: annual growth rate. Please refer to the beginning of the supplemental materials for the definition of the bibliometric indices.

276
277
278
279
280
281
282
283
284
285

• *The most prolific authors*

The top 25 of the most prolific authors included 28 individuals who contributed to at least 7 (7/296, 2.4%) publications (Table MA5). Among these authors, only nine were females (9/28, 32%), and twelve (12/28, 43%) published their last paper more than three years ago. I. Baldi led in total publications (18/296, 6.1%), followed closely by P. Lebailly and E.I. Pukkala, each with 16 publications (16/296, 5.4%). In terms of first authorship, P.K. Mills had a dominant presence, achieving a first author contribution of 100%, followed by P. Petit (85.7%), S. Jouneau (42.9%), and C. Piel (42.9%). D.P. Sandler emerged as the most cited author in this group, with 943 citations, followed by E.I. Pukkala (932) and A. Blair (905). The majority of the articles published by these authors focused on cancer risk (23/28, 82%), with additional research on neurodegenerative disorders (8/28, 29%), particularly Parkinson's disease (7/28, 25%), depression (6/28, 21%), and respiratory conditions (5/28, 18%).

Table MA5. Top 25 of the most prolific authors.

Author	Sex	Publication period	Number of publications	FF (%)	DF1 (%)	DF last (%)	Total citation	Local citation	Average citation per year	h-index	Mean AGR (%)	g-index	m-index	Y-index	Thematic
BALDI I	F	2011-2024	18 (6.1%)	1.61	5.56	27.8	494	33 (6.7%)	35.3	10	3.57	18	0.71	5.10	Cancer, neurobehavior effect, aging, mortality, PD
LEBAILLY P	M	2011-2024	16 (5.4%)	1.45	0.00	37.5	361	32 (8.9%)	25.8	10	12.5	16	0.71	6.00	Cancer, neurobehavior effect, mortality, PD
PUKKALA EI	M	1997-2023	16 (5.4%)	2.66	12.5	62.5	932	25 (2.7%)	34.5	10	-20.4	16	0.36	10.2	Cancer
LYNGE E	F	1988-2024	15 (5.1%)	3.27	20.0	13.3	418	5 (1.2%)	11.3	9	-17.6	15	0.24	3.61	Cancer
TUAL S	F	2017-2022	13 (4.4%)	1.10	15.4	0.00	358	27 (7.5%)	59.7	9	0.00	13	1.12	2.00	Cancer, PD
BLAIR A	M	1981-2017	11 (3.7%)	2.02	18.2	27.3	905	22 (2.4%)	24.5	10	-16.2	11	0.23	3.61	Cancer
SANDLER DP	M	2003-2020	11 (3.7%)	1.27	0.00	0.00	943	17 (1.8%)	52.4	11	-22.2	11	0.50	0.00	Cancer, CVD, injury
BOULANGER M	F	2017-2022	10 (3.4%)	0.93	20.0	0.00	151	11 (7.3%)	25.2	7	5.56	10	0.88	2.00	Cancer, mortality
KJAERHEIM K	F	1999-2023	10 (3.4%)	1.21	20.0	10.0	839	29 (3.5%)	33.6	8	-12.0	10	0.31	2.24	Cancer
DEGANO B	M	2016-2020	9 (3.0%)	1.00	0.00	44.4	123	12 (9.8%)	24.6	5	16.7	9	0.56	4.00	COPD, depression
ELBAZ A	M	2003-2021	9 (3.0%)	1.45	22.2	55.6	549	8 (1.5%)	28.9	8	-26.3	9	0.36	5.39	PD, motor neuron disease
LAPLANTE JJ	M	1998-2019	9 (3.0%)	0.85	0.00	0.00	210	16 (7.6%)	9.55	7	11.4	9	0.26	0.00	Asthma, cancer, COPD, depression
MARCOTULLIO E	F	2017-2022	9 (3.0%)	0.72	0.00	0.00	181	20 (11.0%)	30.2	8	4.17	9	1.00	0.00	Cancer, PD
MARTINSEN JI	M	2009-2023	9 (3.0%)	1.37	0.00	0.00	573	24 (4.2%)	38.2	8	-20.0	9	0.50	0.00	Cancer
BONNETERRE V	M	2020-2024	8 (2.7%)	1.79	12.5	87.5	28	16 (57.1%)	5.60	3	30.0	5	1.00	7.07	AD, cancer, depression, IBD, thyroid disorders
DALPHIN JC	M	1998-2020	8 (2.7%)	0.86	12.5	50.0	195	11 (5.6%)	8.48	6	-4.35	8	0.22	4.12	Asthma, cancer, COPD, depression
FREEMAN LEB	F	2013-2023	8 (2.7%)	0.75	0.00	37.5	380	10 (2.6%)	34.6	7	-22.7	8	0.58	3.00	Cancer, injury
HOPPIN JA	F	2003-2017	8 (2.7%)	0.88	0.00	25.0	803	15 (1.9%)	53.5	8	-20.0	8	0.36	2.00	Cancer, injury
SOUMAGNE TE	M	2016-2020	8 (2.7%)	0.88	12.5	0.00	120	12 (10.0%)	24.0	5	0.00	8	0.56	1.00	COPD, depression
PETIT P	M	2022-2024	7 (2.4%)	1.63	85.7	14.3	21	14 (66.7%)	7.00	3	83.3	4	1.00	6.08	AD, cancer, depression, IBD, thyroid disorders
ALAVANJA MCR	M	2003-2017	7 (2.4%)	0.77	28.6	57.1	752	17 (2.3%)	50.1	7	-13.3	7	0.32	4.47	Cancer, injury
JOUNEAU S	M	2015-2022	7 (2.4%)	0.55	42.9	0.00	69	6 (8.7%)	8.63	4	-25.0	7	0.40	3.00	Asthma, COPD
KOUTROS S	F	2010-2020	7 (2.4%)	0.67	0.00	0.00	369	6 (1.6%)	33.6	6	-9.09	7	0.40	0.00	Cancer
LYNCH CF	M	2003-2023	7 (2.4%)	0.75	0.00	0.00	744	15 (2.0%)	35.4	6	-19.0	7	0.27	0.00	Cancer
MILLS PK	M	2001-2009	7 (2.4%)	3.17	100	0.00	327	6 (1.8%)	36.3	7	-44.4	7	0.29	7.00	Cancer
NORDBY KC	M	2004-2023	7 (2.4%)	0.90	28.6	0.00	183	11 (6.0%)	9.15	5	-20.0	7	0.24	2.00	Cancer
PIEL C	M	2017-2022	7 (2.4%)	0.51	42.9	0.00	200	9 (4.5%)	33.3	6	8.33	7	0.75	3.00	Cancer, PD
TZOURIO C	M	2003-2015	7 (2.4%)	0.95	0.00	42.9	538	7 (1.3%)	41.4	7	-23.1	7	0.32	3.00	PD

286 Note: AD: Alzheimer's disease, AGR: annual growth rate, COPD: chronic obstructive pulmonary disease, CVD: cardiovascular disorders, DF: dominance factor, DF1:
287 dominance factor for being first author, DF last: dominance factor for being last author, F: female, FF: fractionalized frequency, IBD: inflammatory bowel disease, M: male,
288 PD: Parkinson's disease.

289 • *The most active institutions*

290
 291 There were 338 institutions reported. The top 20 most active institutions comprised 24 organizations that contributed eleven or more publications (11/296, 3.7%) (Table MA6).
 292 These institutions were spread across six countries, with 13 from France (13/24, 54%), four from the US (4/24, 17%), three from Finland (3/24, 13%), two from Denmark (2/24,
 293 8%), and one each from Sweden and Norway (1/24, 4%). INSERM (France) led with the highest number of publications (47/296, 15.9%), followed by Karolinska Institutet
 294 (Sweden) (24/296, 8.1%), and the University of Oslo (Norway) (23/296, 7.8%). INSERM was the most cited (1390/9379, 14.8%), followed closely by NIH (US) and NCI (US),
 295 each with 1336 citations (1336/9379, 14.2%). Among the top 20, CHU Bordeaux (35.3%), Université Grenoble Alpes (25.0%), and CHU Grenoble Alpes (25.0%) recorded the
 296 highest mean AGR.

297
 298 **Table MA6.** Top 20 of the most active institutions.

Institution	Country	Publication period	Number of publications (%)	Total citation	Local citation (%)	Average citation per year	Mean AGR	Fractionalized frequency (%)	h-index	g-index	m-index
INSERM	France	2003-2024	47 (15.9%)	1390	47 (3.38%)	63.2	6.19	7.73	20	37	0.91
Karolinska Institutet	Sweden	1987-2024	24 (8.11%)	1214	29 (2.67%)	31.9	-23.2	7.53	15	24	0.40
University of Oslo	Norway	1996-2024	23 (7.77%)	1137	35 (3.93%)	39.2	-17.2	5.71	12	23	0.41
Université de Bordeaux	France	2011-2024	22 (7.43%)	566	34 (6.01%)	40.4	6.43	2.96	13	22	0.93
National Institutes of Health (NIH)	US	1998-2023	21 (7.09%)	1336	29 (2.17%)	51.4	-19.2	4.24	17	21	0.63
National Cancer Institute (NCI)	US	1998-2023	21 (7.09%)	1336	29 (2.17%)	51.4	-19.2	4.24	17	21	0.63
CHU Bordeaux	France	2012-2023	19 (6.42%)	500	33 (6.60%)	41.7	35.3	2.38	11	19	0.85
Unicancer	France	2011-2023	19 (6.42%)	479	32 (6.68%)	36.8	-12.8	2.28	11	19	0.79
Université de Caen-Normandie	France	2011-2024	18 (6.08%)	450	33 (7.33%)	32.1	2.38	2.35	10	18	0.71
Centre François Baclesse	France	2011-2023	16 (5.41%)	429	28 (6.53%)	33.0	-12.8	1.95	10	16	0.71
Finnish Cancer Registry	Finland	2005-2024	16 (5.41%)	864	25 (3.80%)	43.2	-17.3	2.95	10	16	0.39
Assistance Publique – Hôpitaux de Paris (AP-HP)	France	2003-2023	15 (5.07%)	571	7 (1.23%)	27.2	-14.3	2.21	9	15	0.41
University of Copenhagen	Denmark	1997-2024	15 (5.07%)	621	24 (3.86%)	22.2	-14.3	2.57	8	15	0.29
CNRS	France	2011-2024	14 (4.73%)	152	5 (3.29%)	10.9	-7.14	2.32	4	12	0.29
Tampere university	Finland	1980-2024	14 (4.73%)	886	24 (2.71%)	19.7	-11.1	2.92	9	14	0.20
Université de Franche-Comté	France	1998-2020	13 (4.39%)	257	15 (5.84%)	11.2	-7.97	2.74	9	13	0.33
Finnish Institute of Occupational Health	Finland	1997-2020	12 (4.05%)	518	3 (0.58%)	21.6	-16.7	6.83	10	12	0.36
NIEHS	US	2003-2020	12 (4.05%)	1086	17 (1.57%)	60.3	-25.0	2.34	12	12	0.56
Sorbonne Université	France	2003-2019	12 (4.05%)	595	9 (1.51%)	35.0	-20.6	1.68	10	12	0.46
Université Grenoble Alpes (UGA)	France	2019-2024	11 (3.72%)	35	14 (40.0%)	5.83	25.0	1.56	4	5	1.33
CHU Grenoble Alpes	France	2019-2024	11 (3.72%)	35	14 (40.0%)	5.83	25.0	1.56	4	5	1.33
CHU Besançon	France	1998-2019	11 (3.72%)	247	15 (6.07%)	11.2	-9.85	2.58	9	11	0.33
Danish Cancer Society	Denmark	1988-2024	11 (3.72%)	323	0 (0.00%)	8.73	-13.5	4.83	7	11	0.19
University of Iowa	US	2002-2023	11 (3.72%)	829	19 (2.29%)	37.7	-18.2	3.78	8	11	0.35

299
 300 *Note:* AGR: annual growth rate, CNRS: Centre national de la recherche scientifique (French National Centre for Scientific Research), INSERM: Institut national de la santé et
 301 de la recherche médicale (French National Institute of Health and Medical Research), NIEHS: National Institute of Environmental Health Sciences.
 302 Please refer to the beginning of the supplemental materials for the definition of the bibliometric indices.

303 • *The most active funding bodies*

304

305 Funding sources were not reported for 156 (156/296, 52.7%) publications. Among the top 20 of the most active funding bodies, which contributed to 18 or more publications
 306 (18/296, 6.1%), the NIH (US) led with 22 publications (22/296, 7.4%), followed by Mutualité Sociale Agricole (MSA) (France) (19/296, 6.4%), and Medline (US) (18/296,
 307 6.1%). MIAI@Grenoble Alpes (France) (83.3%), the French national cancer institute (INCa) (42.5%), and Ligue Contre Le Cancer (France) (37.5%) demonstrated the highest
 308 mean AGR among this top 20 (Table MA7).

309

310

Table MA7. Top 20 of the most active funding bodies.

Funding bodies	Country	Publication period	Number of publications (%)	Total citation	Local citation (%)	Average citation per year	Mean AGR	Fractionalized frequency (%)	h-index	g-index	m-index
National Institutes Of Health (NIH)	US	1985-2022	22 (7.43%)	840	16 (2.02%)	22.1	-11.0	10.0	18	22	0.45
Mutualité Sociale Agricole (MSA)	France	2009-2023	19 (6.42%)	632	27 (4.27%)	42.1	8.33	4.51	11	19	0.69
Medline	US	1985-2020	18 (6.08%)	664	14 (2.27%)	18.4	-19.9	9.17	15	18	0.38
French National Agency for Research (ANR)	France	2009-2024	15 (5.07%)	423	15 (3.55%)	26.4	-1.56	3.87	6	15	0.38
French national cancer institute (INCa)	France	2015-2022	15 (5.07%)	333	21 (6.31%)	41.6	42.5	3.71	10	15	1.00
National Institute for Occupational Safety & Health (NIOSH)	US	2005-2022	15 (5.07%)	295	13 (4.41%)	16.4	-13.9	8.21	8	15	0.40
ANSES	France	2010-2022	14 (4.73%)	278	25 (8.99%)	21.4	0.00	3.26	9	14	0.60
Centre François Baclesse	France	2017-2022	12 (4.05%)	291	26 (8.93%)	48.5	0.00	1.31	9	12	1.12
Ligue Contre Le Cancer	France	2017-2022	12 (4.05%)	272	25 (9.19%)	45.3	37.5	1.61	9	12	1.12
Association Pour La Recherche Sur Le Cancer (ARC)	France	2017-2022	11 (3.72%)	268	20 (7.46%)	44.7	3.33	1.17	9	11	1.12
Fondation De France	France	2017-2022	11 (3.72%)	268	20 (7.46%)	44.7	3.33	1.17	9	11	1.12
French National Institute of Agricultural Medicine (INMA)	France	2017-2022	11 (3.72%)	270	26 (9.63%)	45.0	2.78	1.18	9	11	1.12
French Ministry of Higher education and scientific research	France	2009-2024	10 (3.38%)	433	16 (3.70%)	27.1	-25.0	1.47	8	10	0.50
French National Office for Water and Aquatic Environments (Onema)	France	2017-2022	9 (3.04%)	189	25 (13.2%)	31.5	5.56	1.69	6	9	0.75
Miai@Grenoble Alpes	France	2022-2024	7 (2.36%)	21	14 (66.7%)	7.00	83.3	2.83	3	4	1.00
National Institute Of Environmental Health Sciences (NIEHS)	US	2010-2020	7 (2.36%)	397	0 (0.00%)	36.1	-22.7	2.62	6	7	0.40
Novartis Pharma	Switzerland	2018-2022	7 (2.36%)	33	4 (12.1%)	6.60	13.3	2.62	4	5	0.57
National Cancer Institute (NCI)	US	2014-2023	6 (2.03%)	316	1 (0.32%)	31.6	-20.0	1.70	4	6	0.36
Nordic Cancer Union	Nordic countries	2009-2023	6 (2.03%)	514	21 (4.09%)	34.3	-20.0	3.27	4	6	0.25
Centers for Disease Control and Prevention (CDC)	US	2009-2022	5 (1.69%)	32	0 (0.00%)	2.29	-21.4	1.54	3	5	0.19

311

312

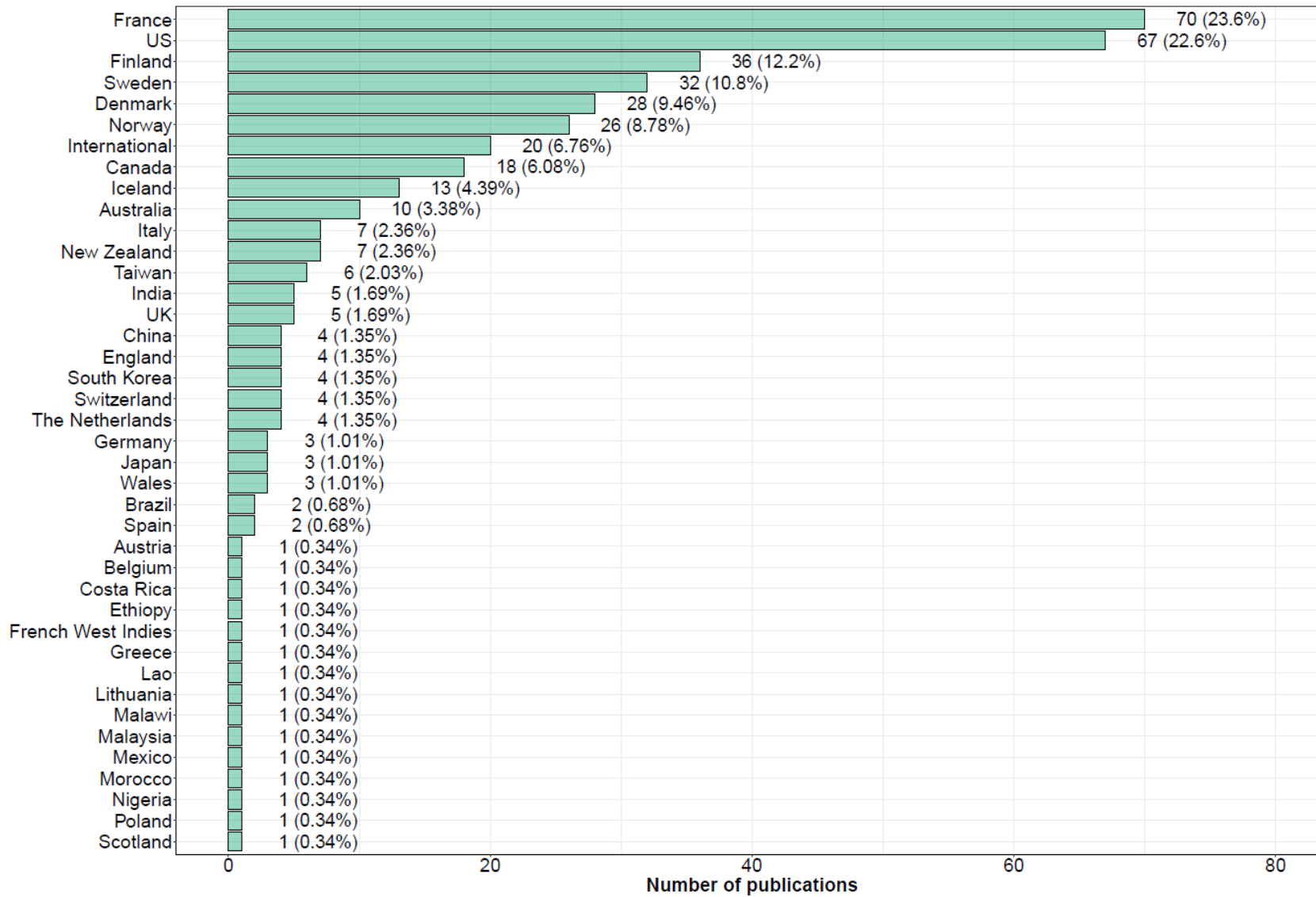
313

Note: AGR: annual growth rate, ANSES: French Agency for Food, Environmental and Occupational Health & Safety.

Please refer to the beginning of the supplemental materials for the definition of the bibliometric indices.

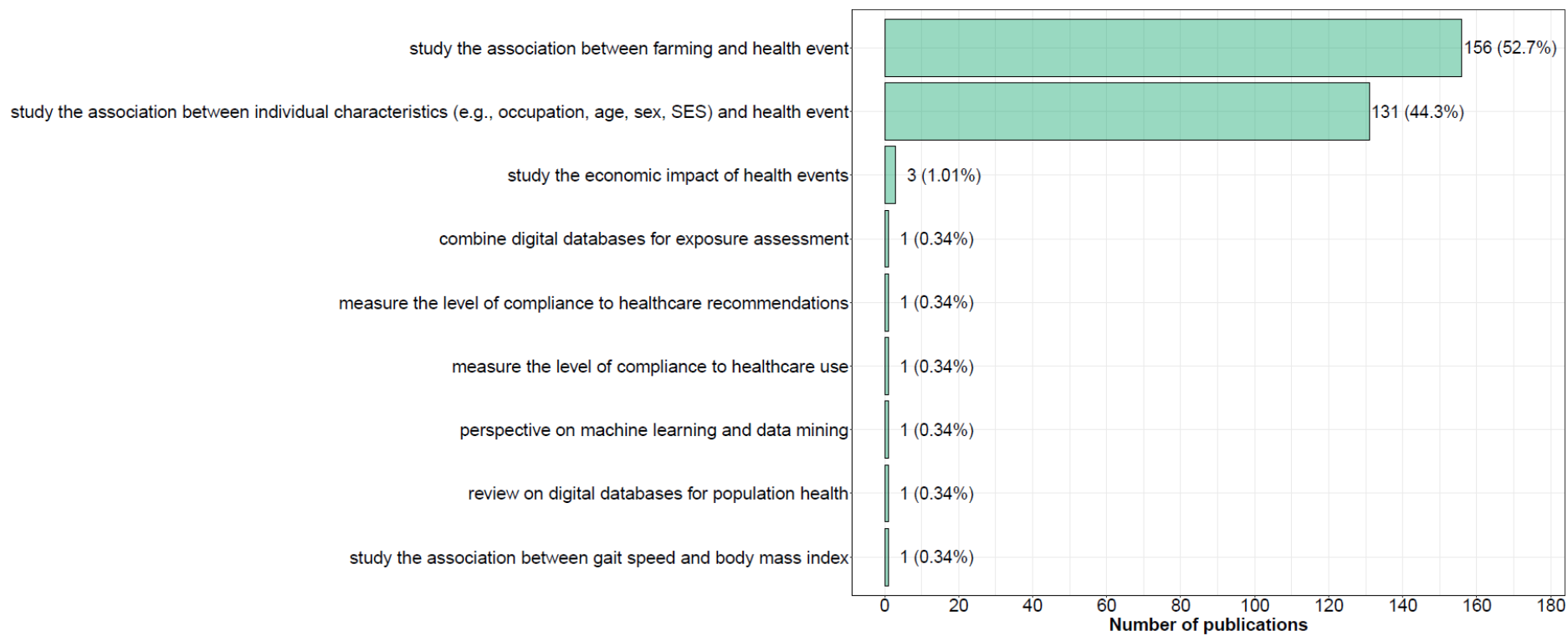
314 **Multimedia appendix 2.** Publications included and analyzed.
315
316 Please refer to the MS Excel file entitled “Multimedia appendix 2”.

317 **Figure S5.** Number of publications per country.



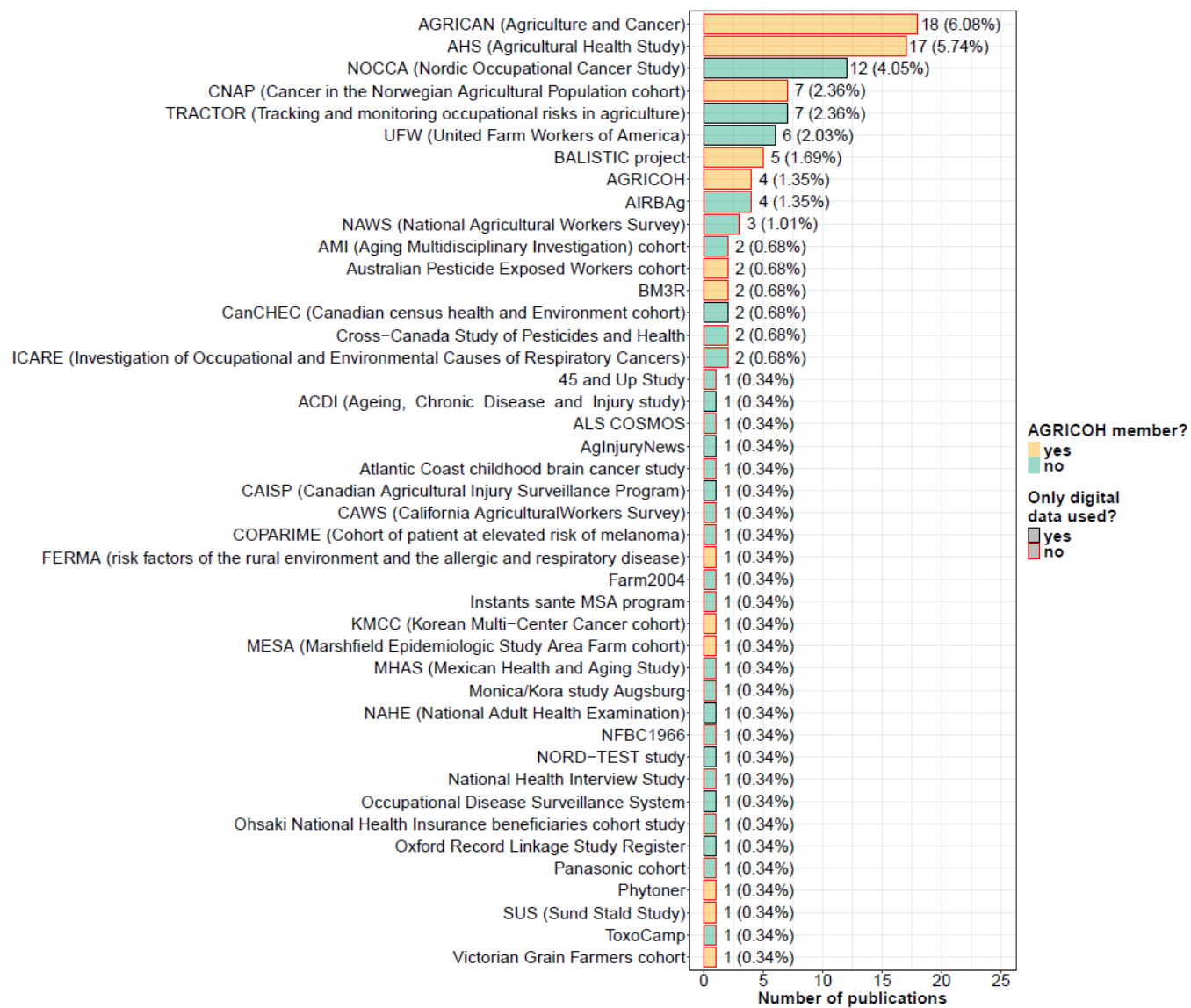
318

319 **Figure S6.** Main goal addressed by study.



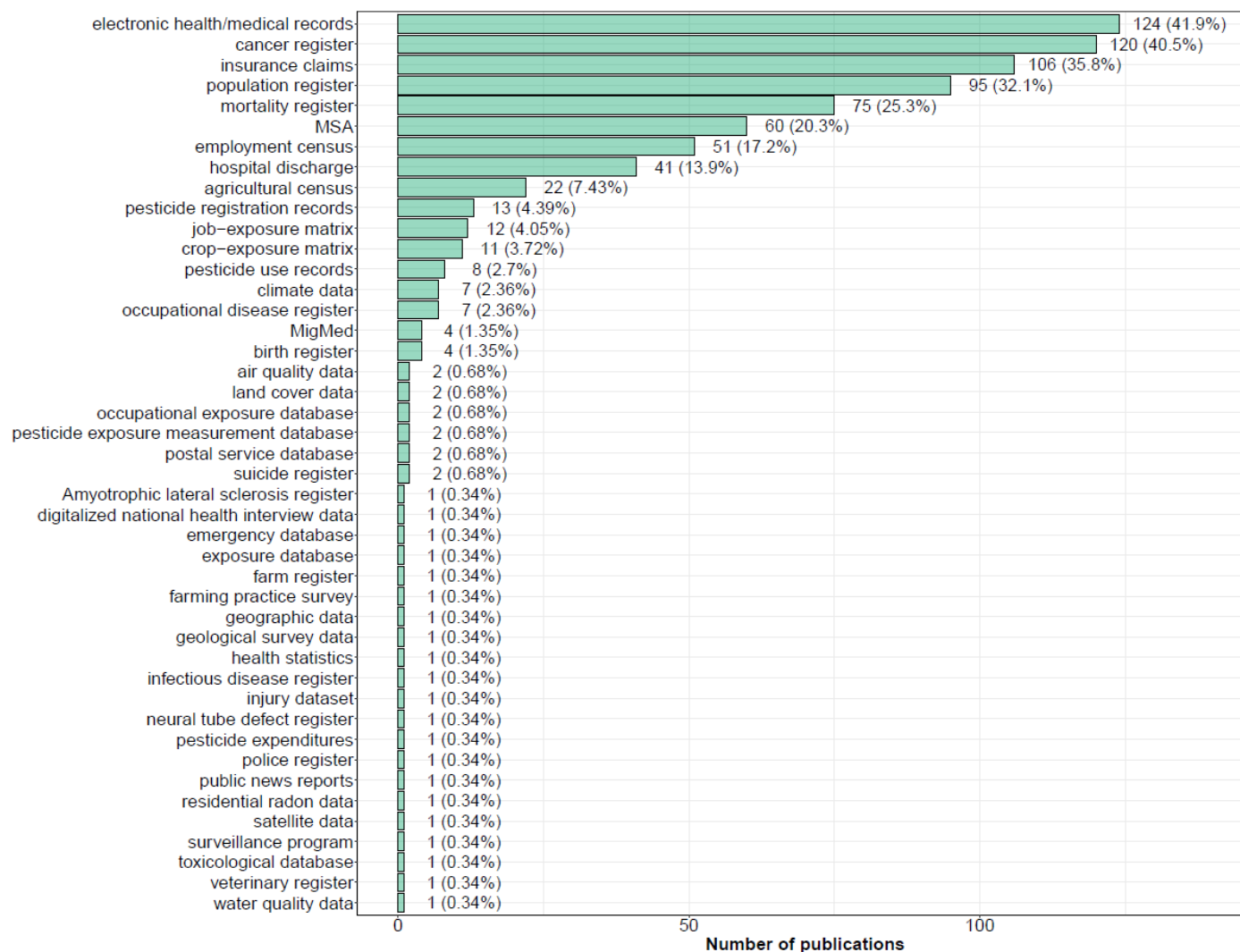
320

321 **Figure S7.** Number of publications per cohort.



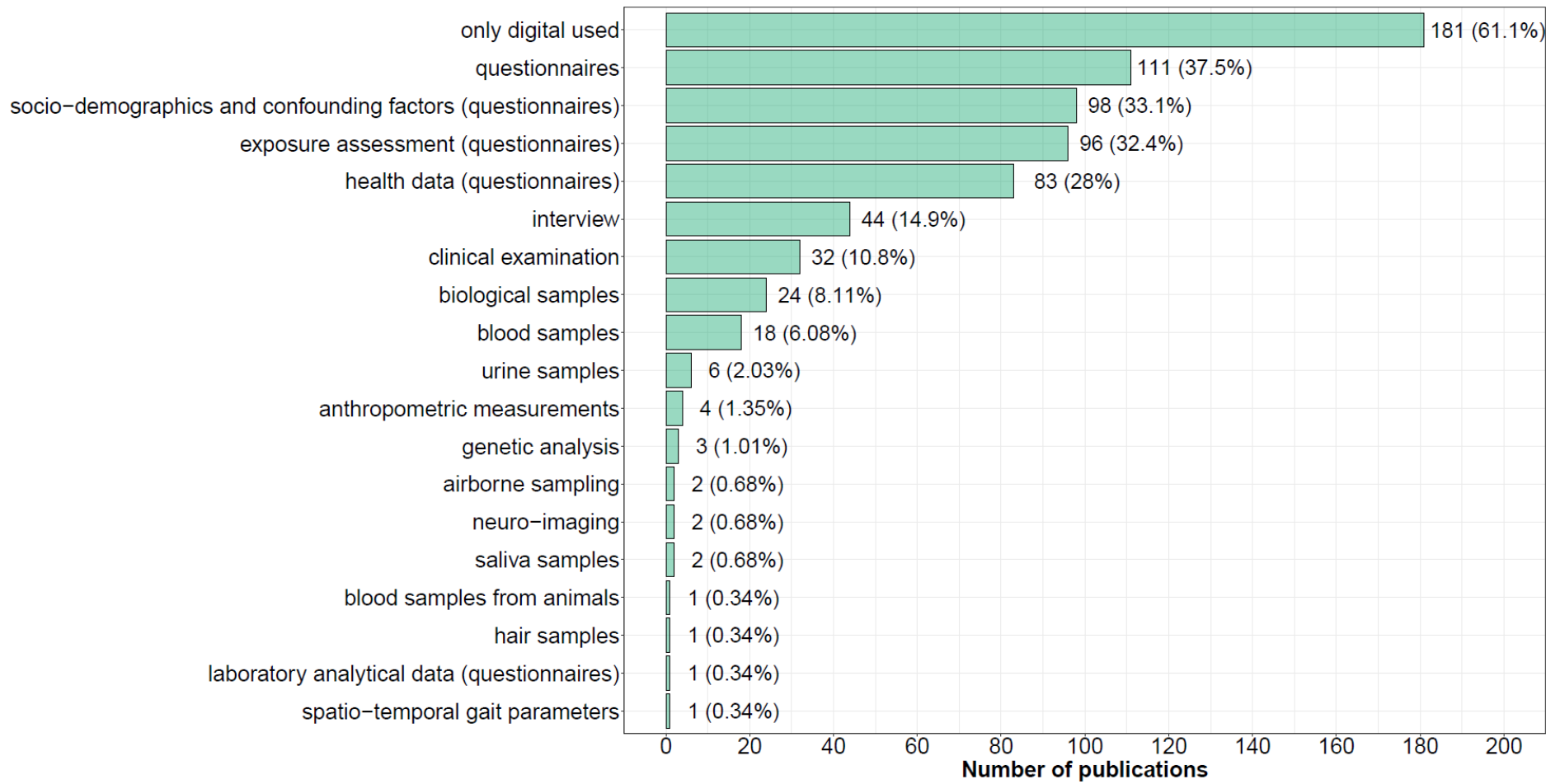
322

323 **Figure S8.** Type of digital data used.



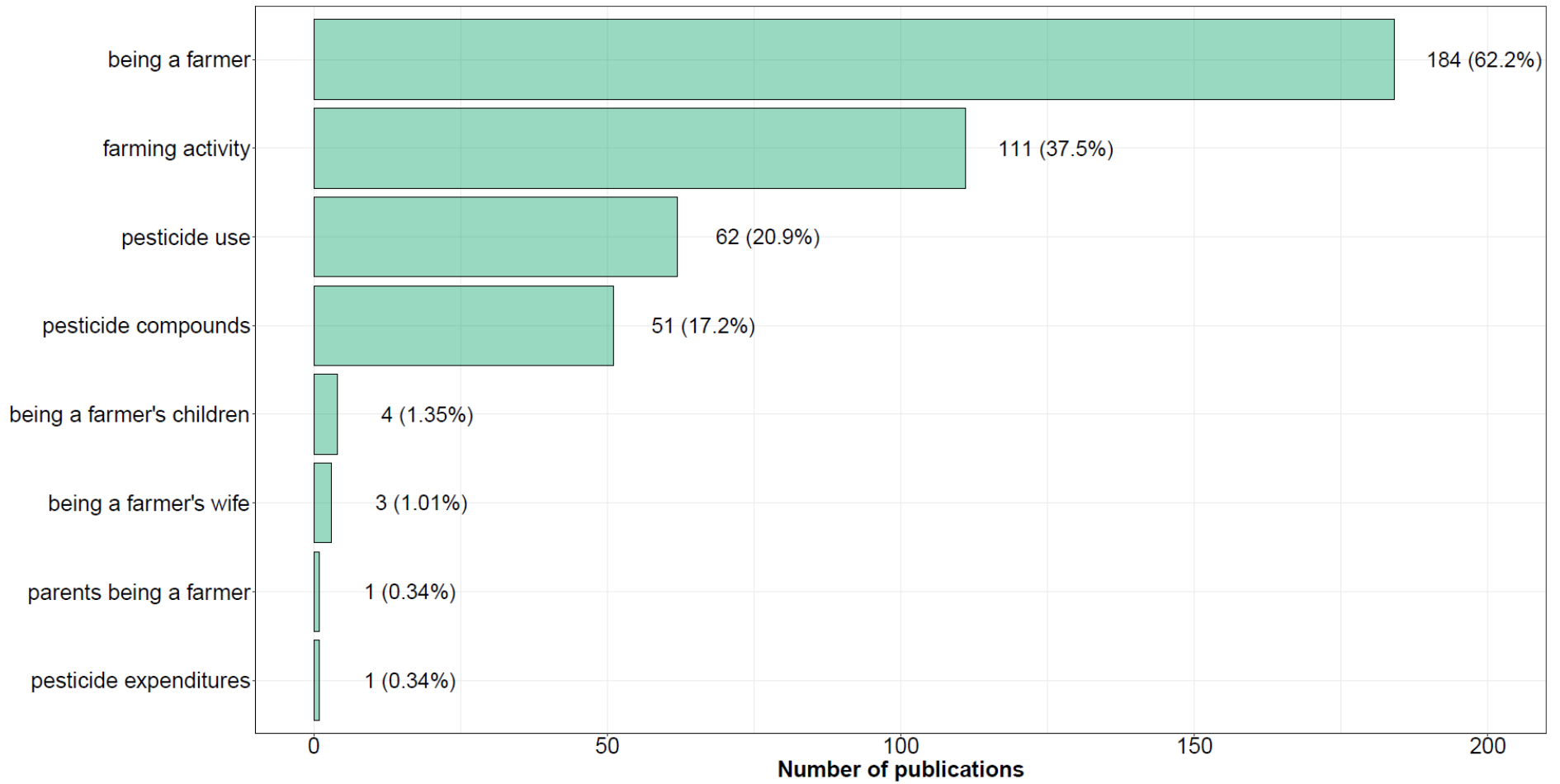
324

325 **Figure S9.** Active data used.



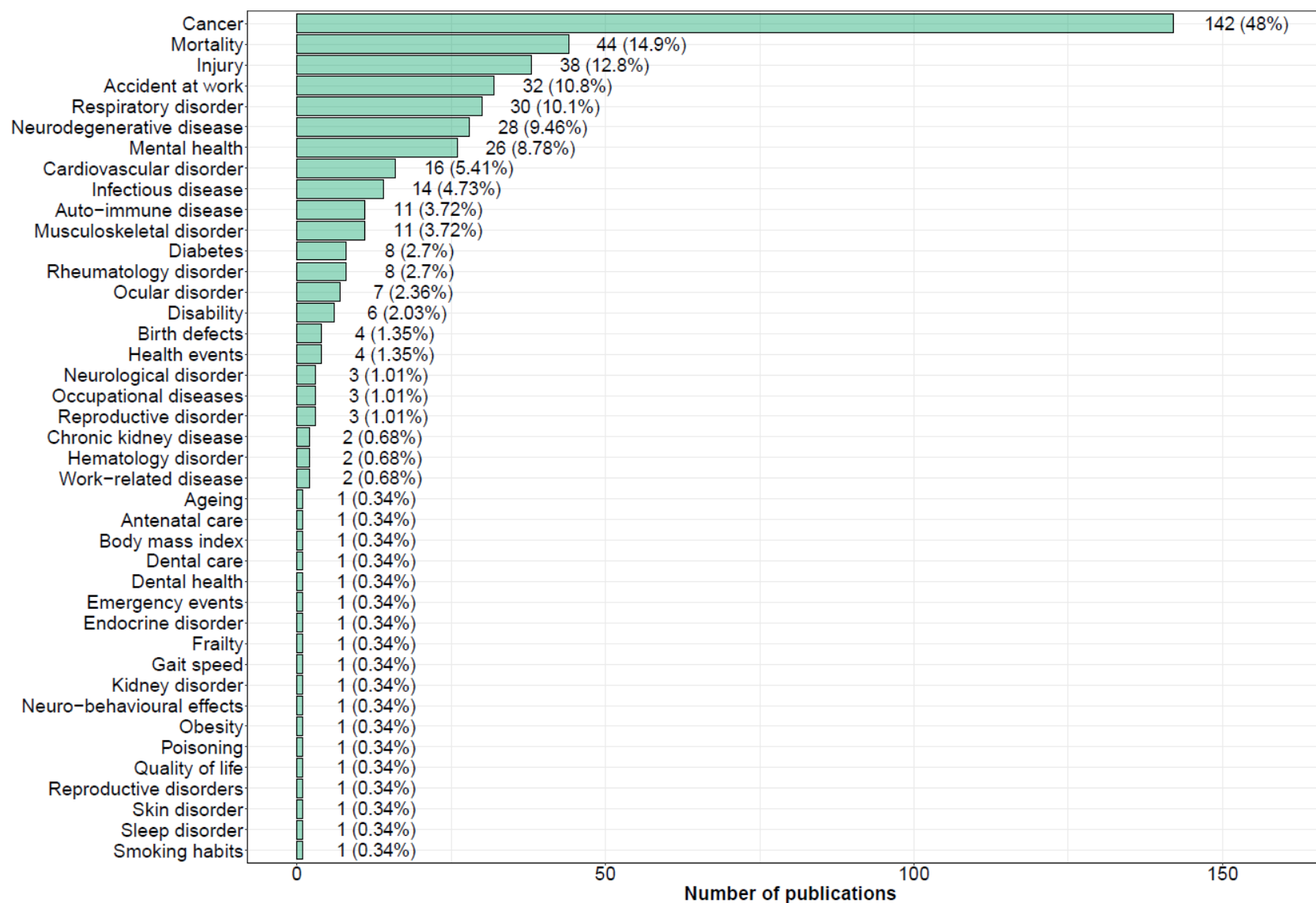
326

327 **Figure S10.** Most frequent farming exposure proxy.



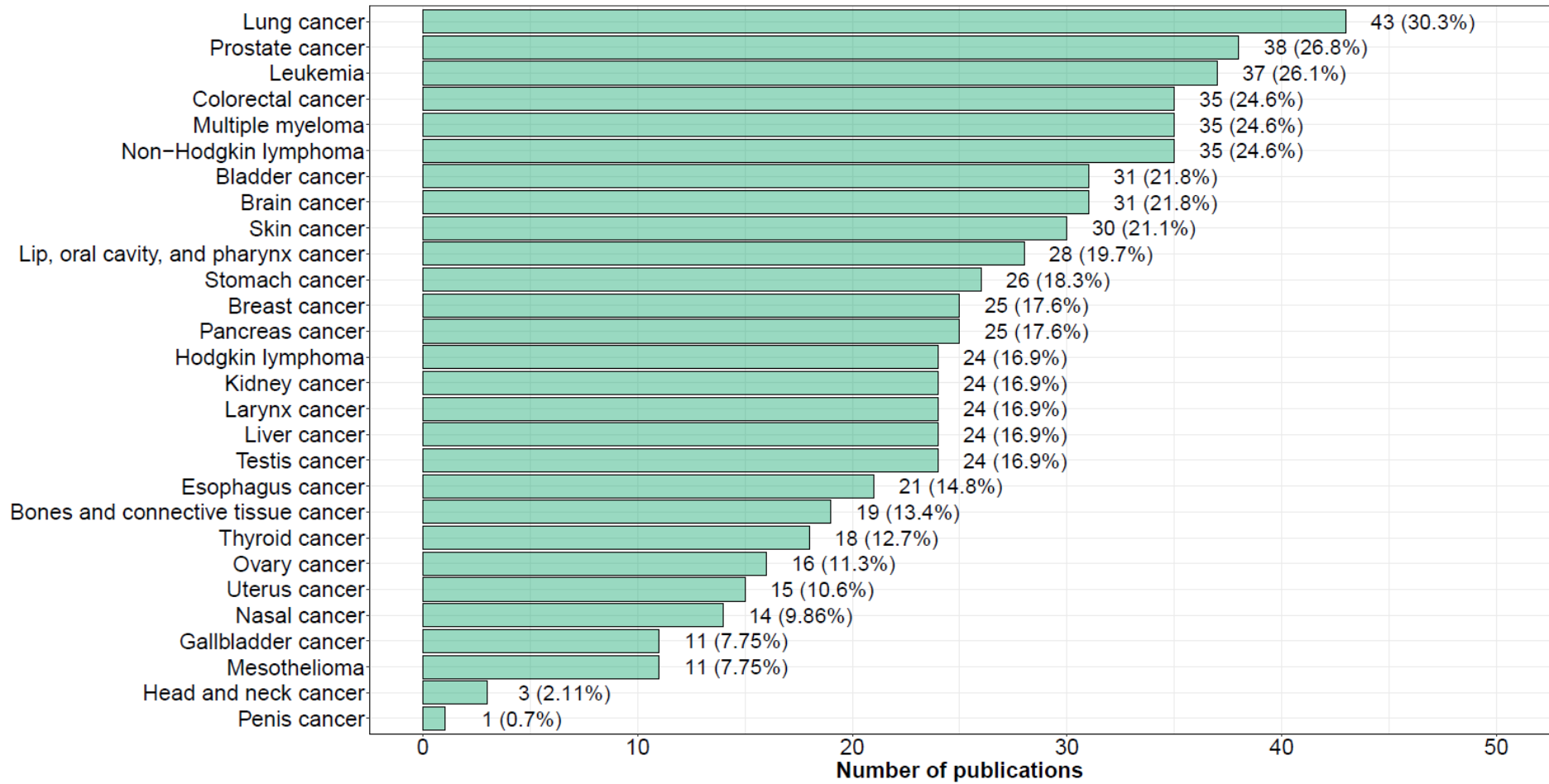
328

329 **Figure S11.** Most frequent health events studied.



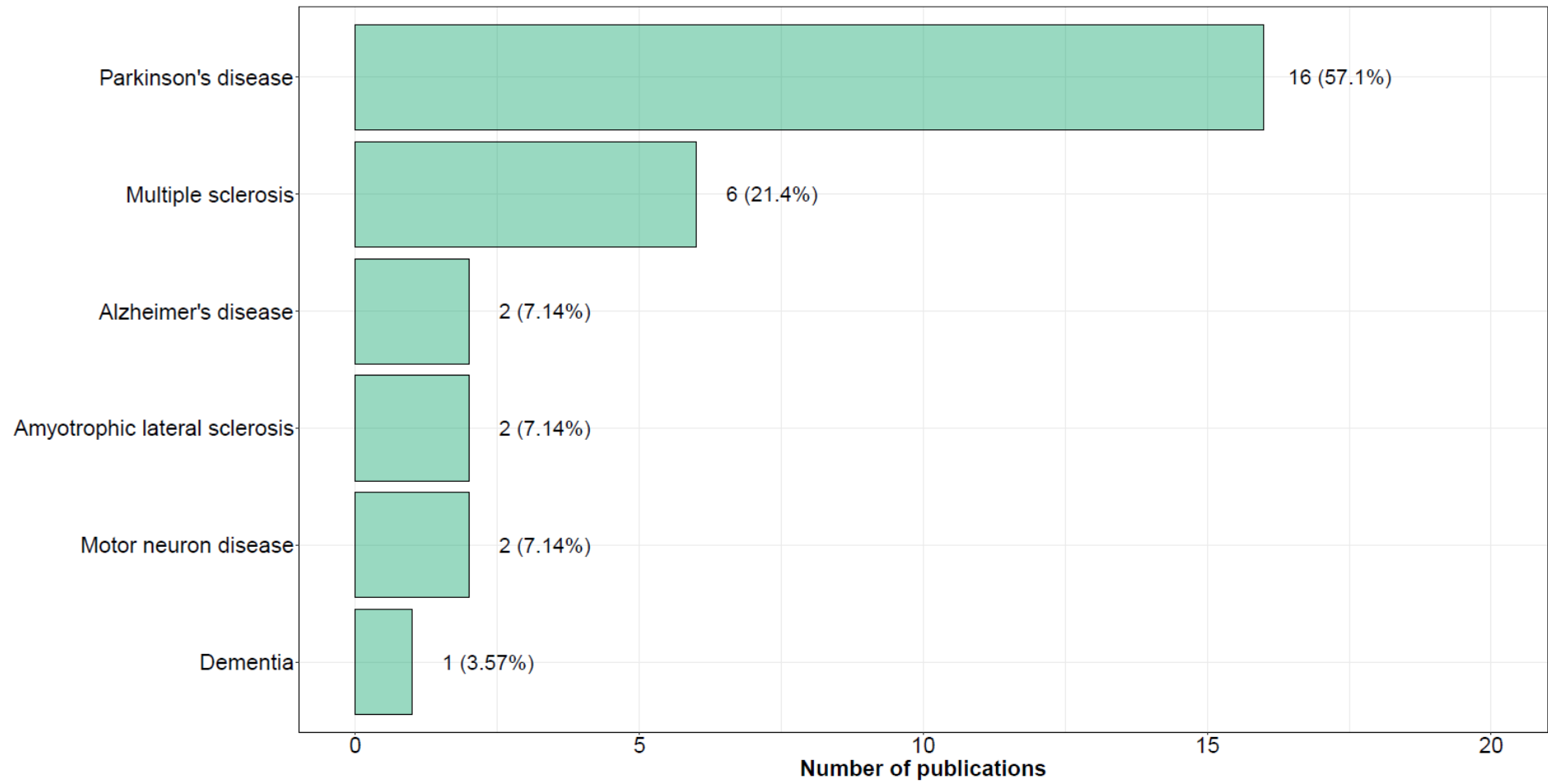
330

331 **Figure S12.** Most frequent cancer studied.



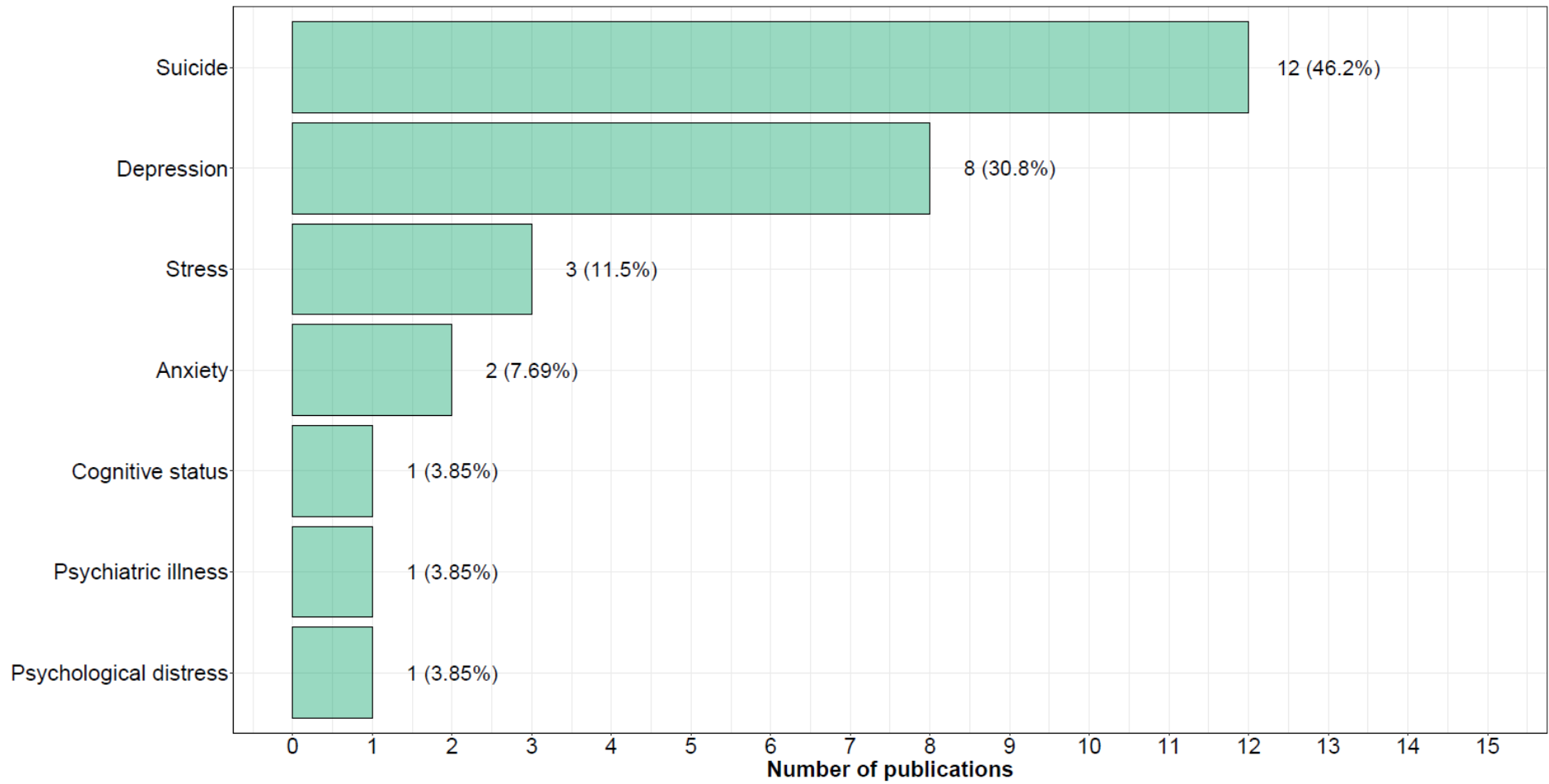
332

333 **Figure S13.** Most frequent neurodegenerative disease studied.



334

335 **Figure S14.** Most frequent mental health issue studied.



336

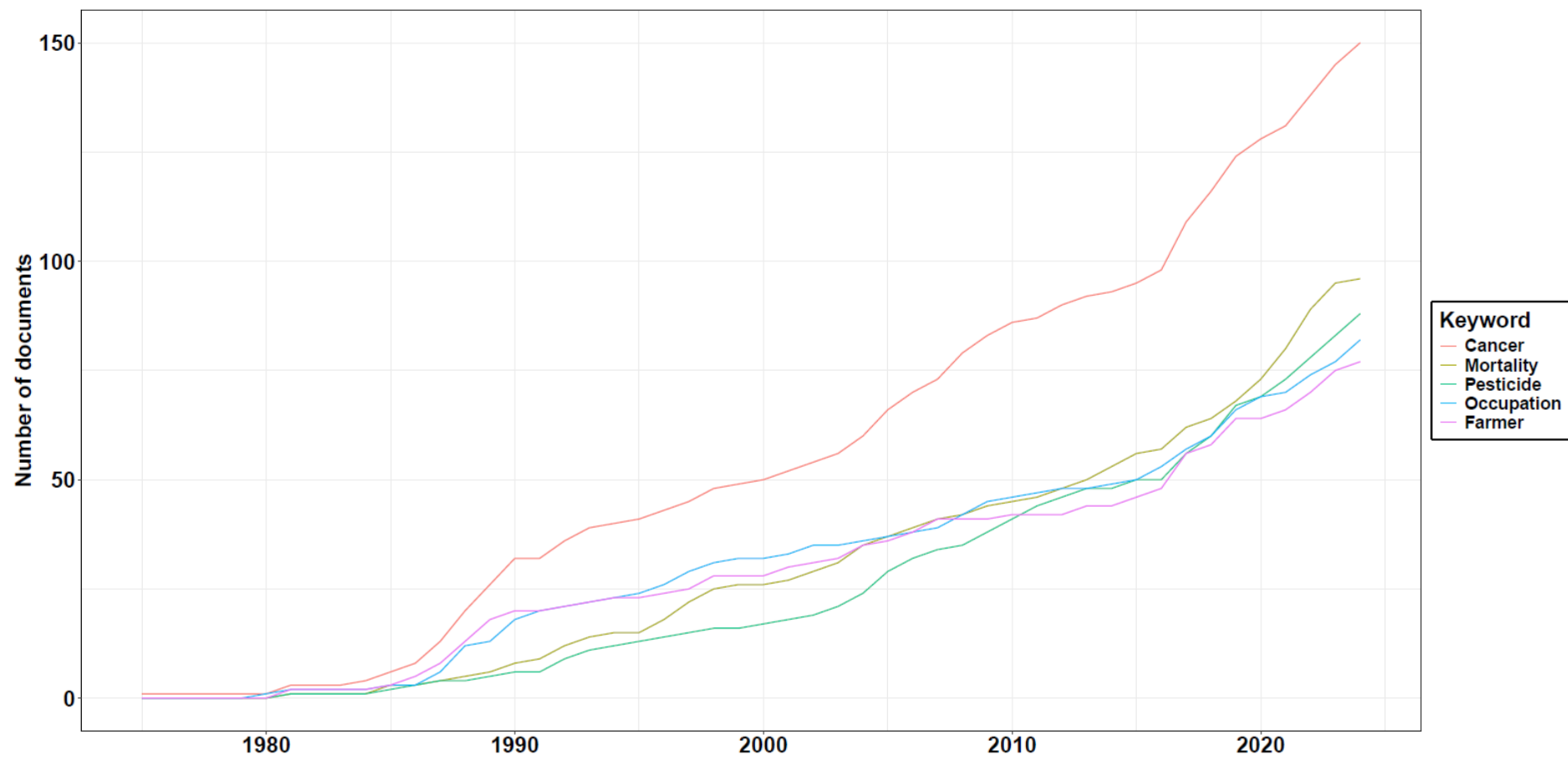
337 **Table S7.** Top 50 of the most frequent keywords.

Keyword	First appearance	Last appearance	Number of publications	Total citation	Local citation	Average citation per year	Mean AGR	h-index	g-index	m-index
Cancer	1975	2024	150 (50.7%)	5766	137 (2.74%)	115	13.5	43	70	0.860
Mortality	1981	2024	96 (32.4%)	3097	75 (2.81%)	70.4	8.15	34	53	0.773
Pesticide	1981	2024	88 (29.7%)	3569	81 (2.44%)	81.1	-7.46	36	58	0.818
Occupation	1980	2024	82 (27.7%)	2959	55 (2.33%)	65.8	14.2	31	53	0.689
Farmer	1981	2024	77 (26.0%)	2856	91 (3.25%)	64.9	6.36	34	52	0.773
Agriculture	1980	2024	74 (25.0%)	1701	55 (3.85%)	37.8	-9.96	26	39	0.578
Epidemiology	1991	2024	57 (19.3%)	1396	17 (1.47%)	41.1	23.1	21	37	0.618
Exposure	1992	2024	57 (19.3%)	1713	27 (1.76%)	61.2	-2.73	21	41	0.636
Lymphohematopoietic cancer	1981	2023	54 (18.2%)	2906	66 (2.69%)	67.6	-13.8	31	53	0.705
Respiratory cancer	1987	2024	51 (17.2%)	2566	67 (3.29%)	67.5	-16.1	28	50	0.737
Risk	1992	2024	51 (17.2%)	1909	27 (1.53%)	57.8	6.01	25	43	0.758
Male genital organ cancer	1987	2024	47 (15.9%)	2865	75 (3.12%)	75.4	-12.5	29	47	0.763
Risk factor	1997	2024	44 (14.9%)	818	17 (2.08%)	29.2	-19.0	17	27	0.607
Digestive organ cancer	1987	2023	41 (13.9%)	2331	58 (3.09%)	63.0	-11.3	26	41	0.684
Occupational exposure	1992	2024	41 (13.9%)	789	39 (4.94%)	23.9	2.01	17	27	0.515
Worker	1992	2024	41 (13.9%)	1243	32 (2.81%)	37.7	-4.85	19	35	0.576
Injury	1991	2024	40 (13.5%)	734	18 (2.69%)	21.6	-5.44	15	26	0.441
Cohort	1992	2024	39 (13.2%)	879	40 (4.55%)	26.6	1.41	16	29	0.485
Incidence	1985	2023	34 (11.5%)	1187	29 (3.21%)	30.4	-13.7	20	34	0.500
Neurodegenerative disease	2000	2024	33 (11.1%)	1216	13 (1.12%)	55.3	-10.0	17	33	0.680
Urinary tract cancer	1987	2023	33 (11.1%)	1975	47 (3.03%)	53.4	-9.59	21	33	0.553
Workplace accident	1987	2024	32 (10.8%)	727	17 (2.76%)	20.8	-6.58	15	26	0.395
Respiratory disorder	1986	2024	32 (10.8%)	901	21 (2.46%)	23.1	-15.2	15	30	0.385
US	1994	2024	32 (10.8%)	1167	14 (1.26%)	40.2	-25.0	18	32	0.581
Brain cancer	1981	2024	31 (10.5%)	1858	61 (3.58%)	42.2	-10.6	23	31	0.523
Health	1997	2023	31 (10.5%)	463	25 (5.45%)	17.1	-12.0	12	21	0.429
Lip, oral cavity, and pharynx cancer	1975	2024	30 (10.1%)	1827	55 (3.93%)	36.5	-14.8	21	30	0.420
Male	1985	2023	30 (10.1%)	1150	29 (2.86%)	29.5	-15.4	19	30	0.475
Skin cancer	1975	2023	30 (10.1%)	1999	44 (2.80%)	40.8	-22.8	22	30	0.440
Prevalence	1998	2024	28 (9.46%)	438	17 (3.88%)	16.2	-14.8	12	20	0.444
Socio-economic status	1980	2024	27 (9.12%)	685	7 (1.33%)	15.2	-24.4	16	26	0.356
Animal farming	1975	2023	26 (8.78%)	883	29 (3.28%)	18.0	-20.1	15	26	0.300
Breast cancer	1987	2024	26 (8.78%)	1628	48 (3.63%)	42.8	-18.7	20	26	0.526
Farming activity	1985	2023	25 (8.45%)	925	35 (3.82%)	23.7	-10.7	14	25	0.350
Mental health	1991	2024	25 (8.45%)	512	6 (1.22%)	15.1	-28.4	12	22	0.353
Asthma	1986	2024	24 (8.11%)	743	5 (0.72%)	19.1	-12.4	13	24	0.333
Farm	1992	2024	24 (8.11%)	516	8 (1.75%)	15.6	-6.57	10	22	0.303
Herbicide	1984	2024	22 (7.43%)	1349	26 (1.93%)	32.9	-12.6	15	22	0.366
Mesenchymal cancer	1984	2023	20 (6.76%)	1466	35 (3.07%)	36.6	-18.1	14	20	0.341
COPD	1987	2024	19 (6.42%)	298	14 (5.60%)	11.0	-1.32	8	17	0.211
Environmental exposure	1999	2024	19 (6.42%)	688	12 (3.20%)	31.3	-17.9	11	19	0.423
Female	1980	2024	19 (6.42%)	454	7 (1.54%)	10.1	-25.6	11	19	0.244
Female genital organ cancer	1987	2024	19 (6.42%)	1083	47 (4.34%)	28.5	-21.1	14	19	0.368
Insecticide	1992	2023	19 (6.42%)	1218	33 (2.71%)	38.1	-21.9	14	19	0.424
Parkinson's disease	2000	2021	19 (6.42%)	972	12 (1.32%)	51.2	-19.7	16	19	0.640
Pesticide applicator	1987	2024	19 (6.42%)	984	32 (3.57%)	25.9	-30.7	14	19	0.368
Smoking	1989	2022	19 (6.42%)	586	14 (2.66%)	17.2	-18.6	12	19	0.333
Assessment	2003	2024	18 (6.08%)	293	0 (0.00%)	13.3	-18.9	10	17	0.455
CVD	1987	2023	17 (5.74%)	623	3 (0.52%)	18.3	-18.9	11	17	0.289

Disease	1997	2023	17 (5.74%)	385	2 (0.58%)	14.3	-11.1	8	17	0.286
Follow-up	1997	2024	17 (5.74%)	562	5 (0.96%)	20.1	-23.8	10	17	0.357
Surveillance	1988	2024	17 (5.74%)	291	13 (5.04%)	7.86	-17.6	8	17	0.216

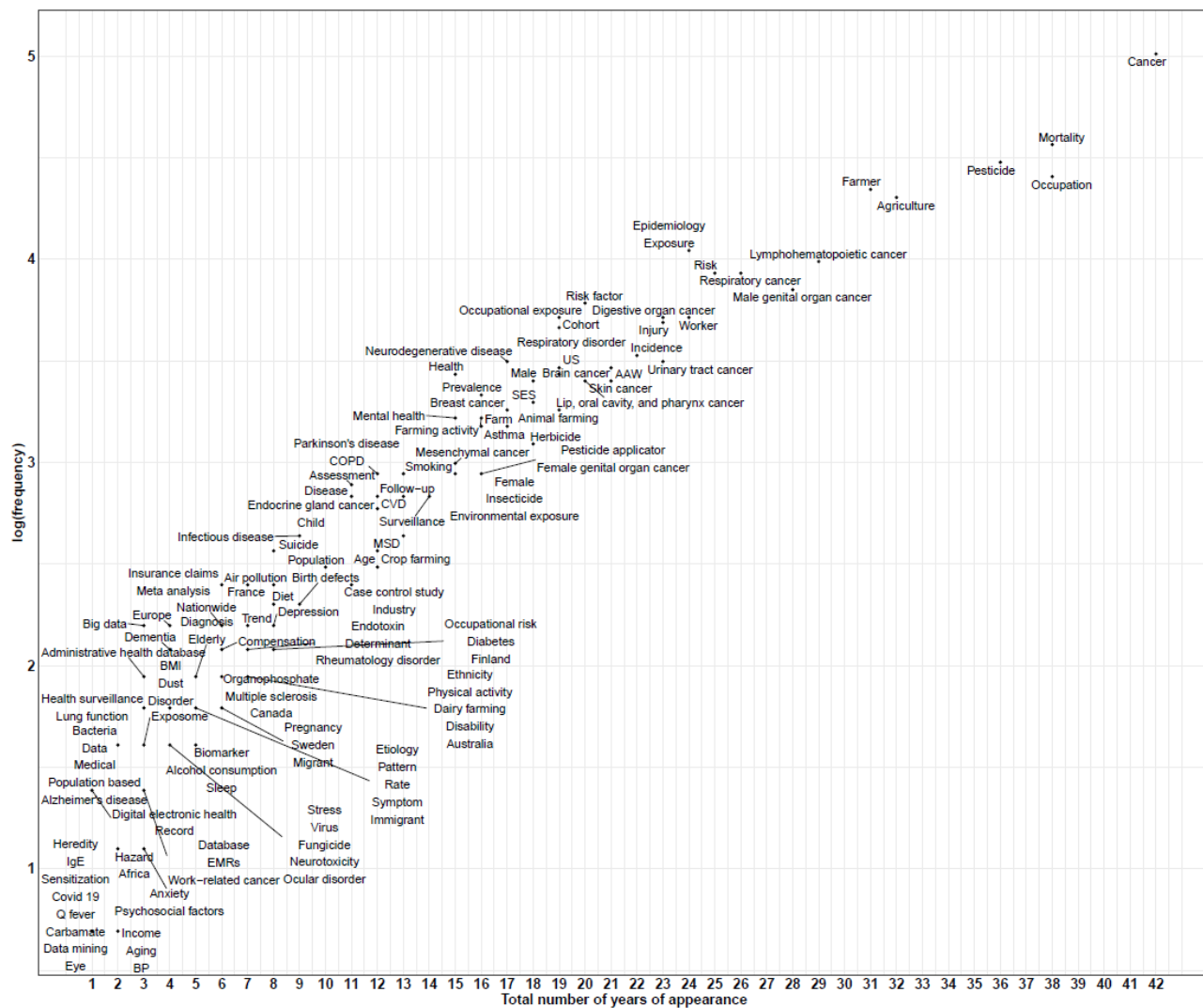
338 *Note:* AGR: annual growth rate, COPD: chronic obstructive pulmonary disease, CVD: cardiovascular disorder. Please refer to the beginning of the supplemental materials for
339 the definition of the bibliometric indices.

340 **Figure S15.** Top 5 of the most frequent keywords – growth trends.



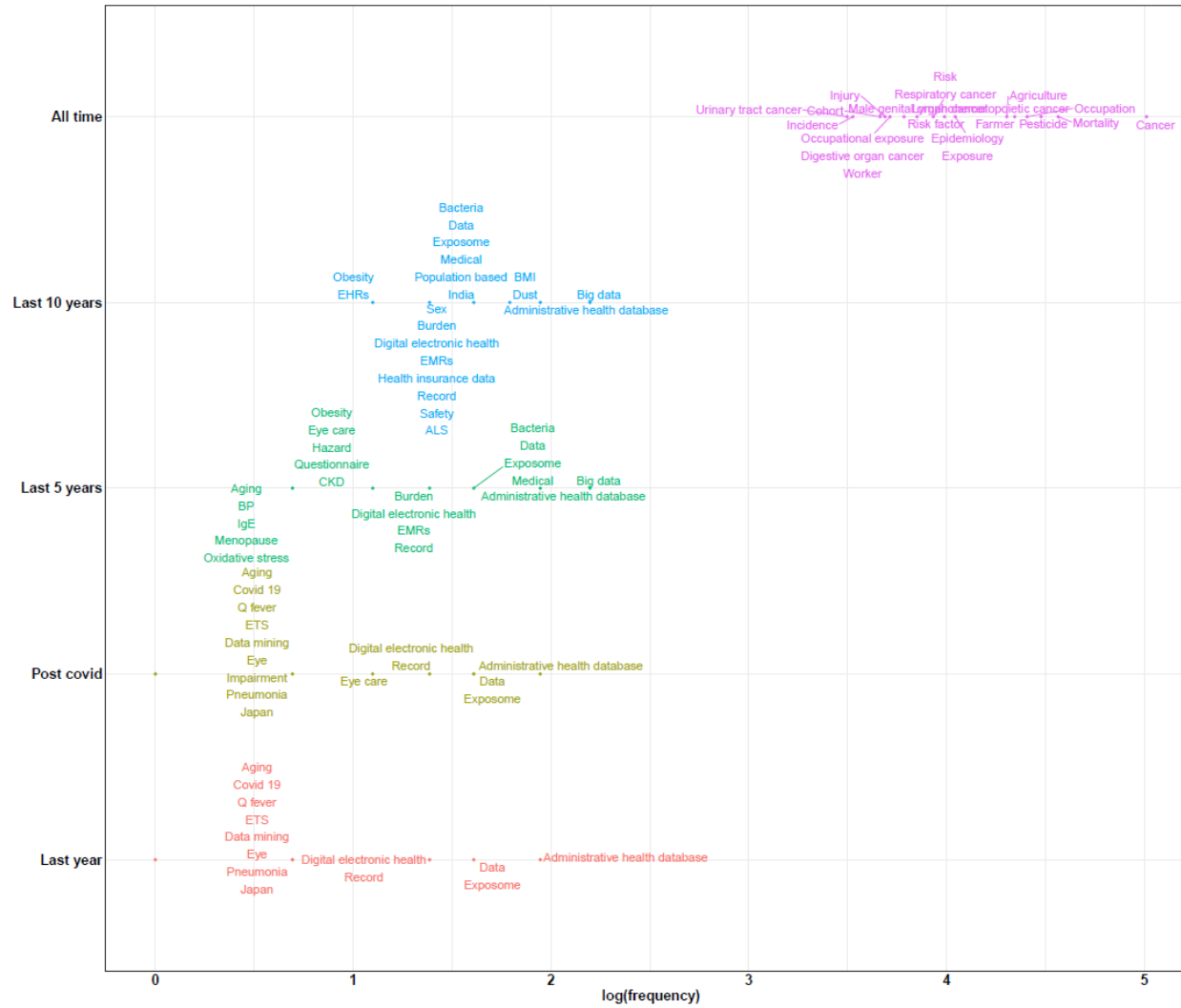
341

342 **Figure S16.** Top ten of the most popular keywords by total number of years of appearance.
 343 AAW: workplace accident; ALS: amyotrophic lateral sclerosis, BMI: body mass index, BP: blood pressure, COPD: chronic obstructive pulmonary disease, CVD: cardiovascular
 344 disorder, EMR: electronic medical record, IgE: immunoglobulin E, MSD: musculoskeletal disorder, SES: socio-economic status, US: United States of America.



345

346 **Figure S17.** Top 10 of the most frequent keywords by time period.
 347 BMI: body mass index, EMR: electronic medical record, ETS: environmental tobacco smoking.



348

349

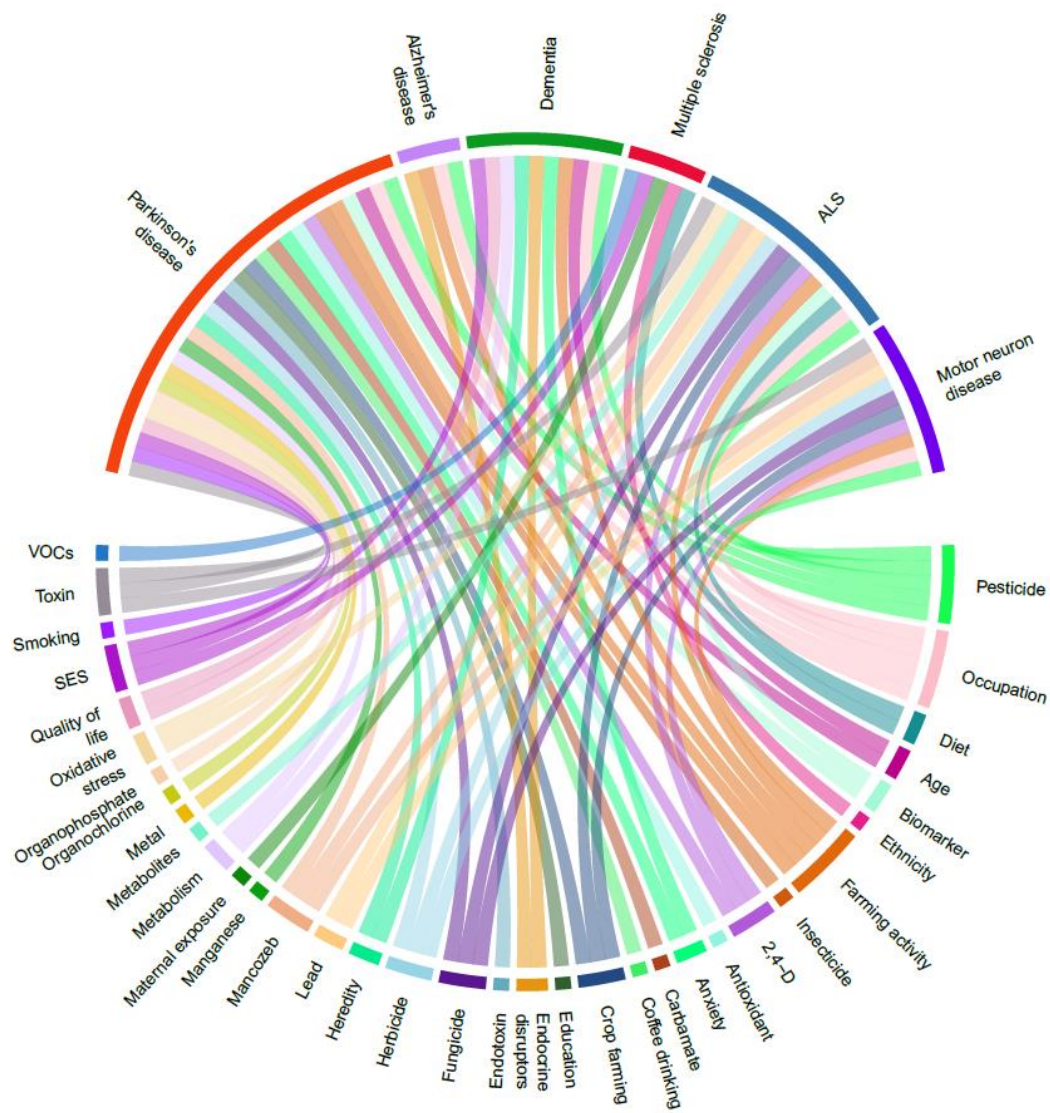
350 **Multimedia appendix 3.** Co-occurrence between exposome-related keywords and health event-related keywords.
351
352 Please refer to the MS Excel file entitled “Multimedia appendix 3”.

353 **Figure S18.** Keyword co-occurrence network between potential risk factor and mental health disorder keywords.
354 BMI: body mass index, SES: socio-economic status. Larger the node size, higher the number of co-occurrence. Red nodes refer to health events while blue nodes refer to
355 potential risk factors.

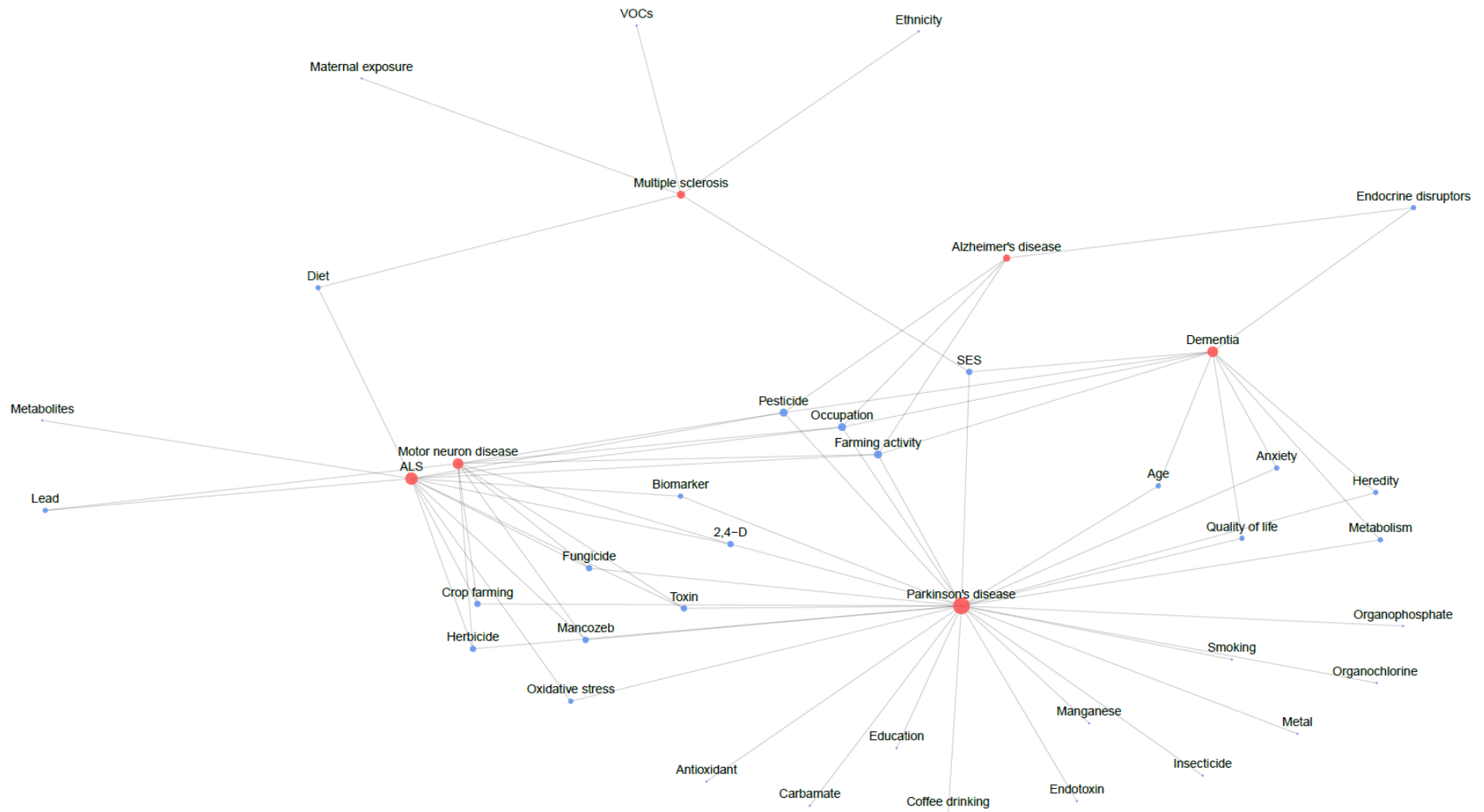


356

357 **Figure S19.** Chord diagram of keyword co-occurrence between potential risk factor and neurodegenerative disease keywords.
 358 ALS: amyotrophic lateral sclerosis, SES: socio-economic status, VOC: volatile organic compound.

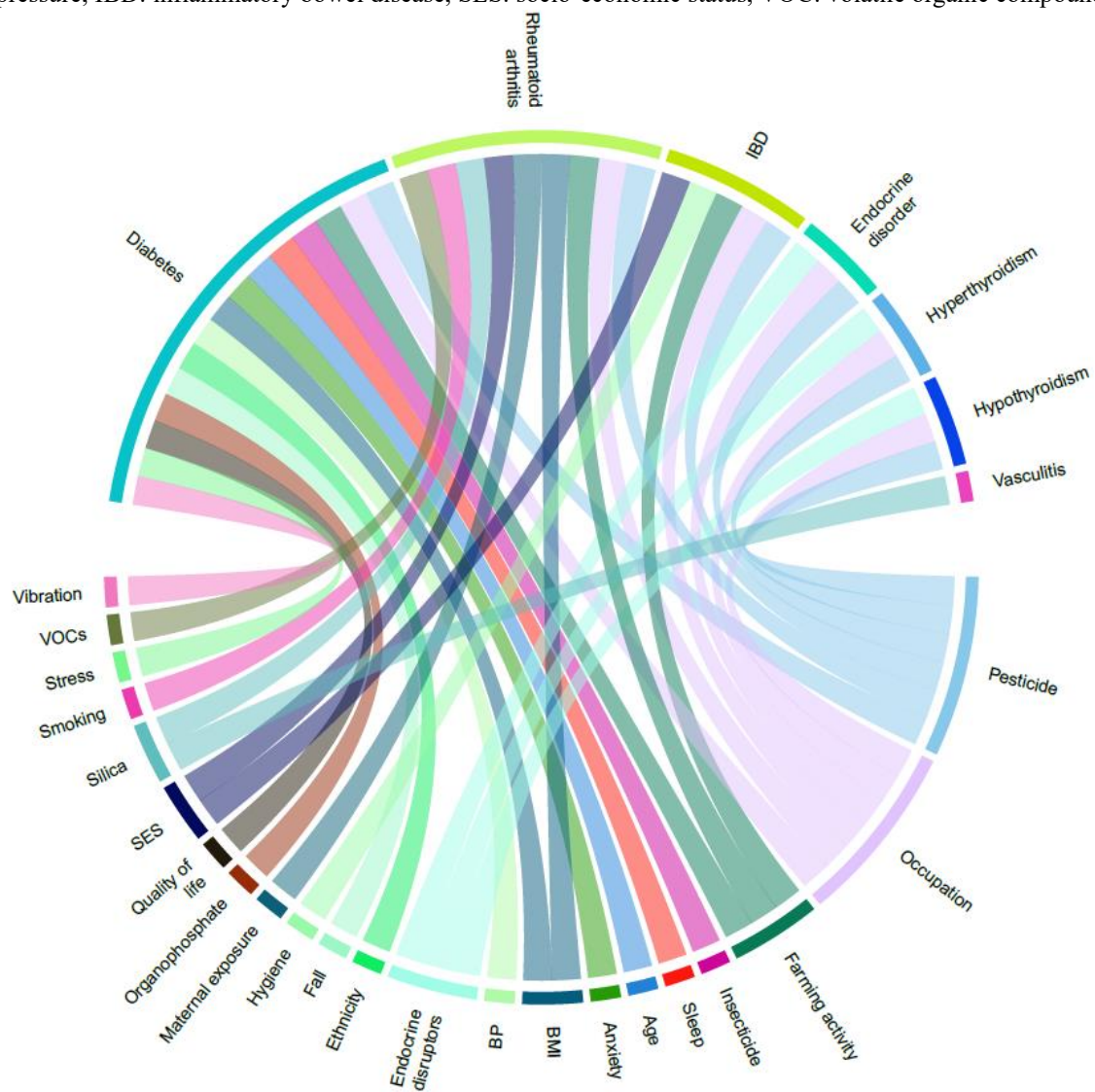


360 **Figure S20.** Keyword co-occurrence network between potential risk factor and neurodegenerative disease keywords.
 361 ALS: amyotrophic lateral sclerosis, SES: socio-economic status, VOC: volatile organic compound. Larger the node size, higher the number of co-occurrence. Red nodes refer
 362 to health events while blue nodes refer to potential risk factors.

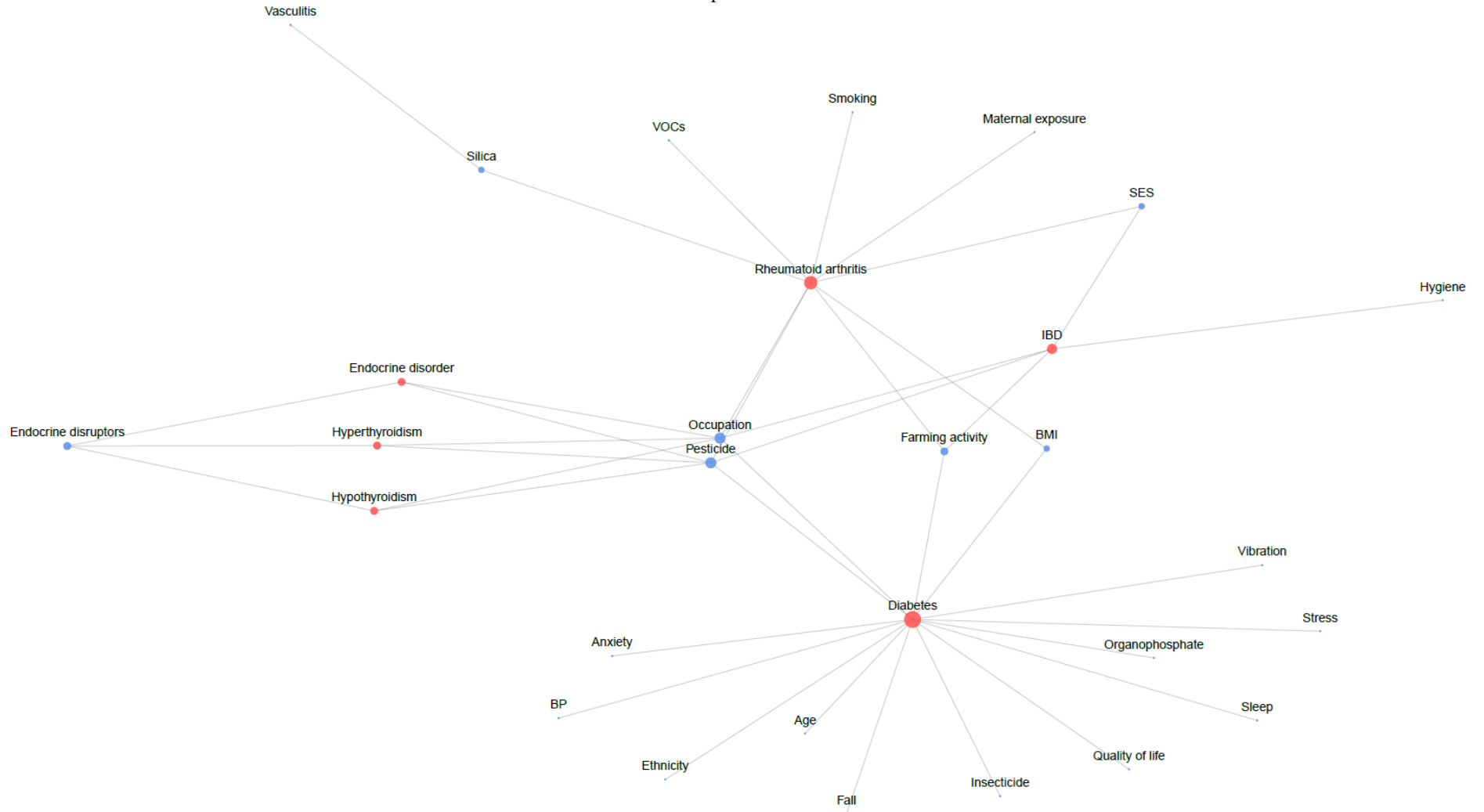


363
 364

365 **Figure S21.** Chord diagram of keyword co-occurrence between potential risk factor and autoimmune disease keywords.
 366 BMI: body mass index, BP: blood pressure, IBD: inflammatory bowel disease, SES: socio-economic status, VOC: volatile organic compound.

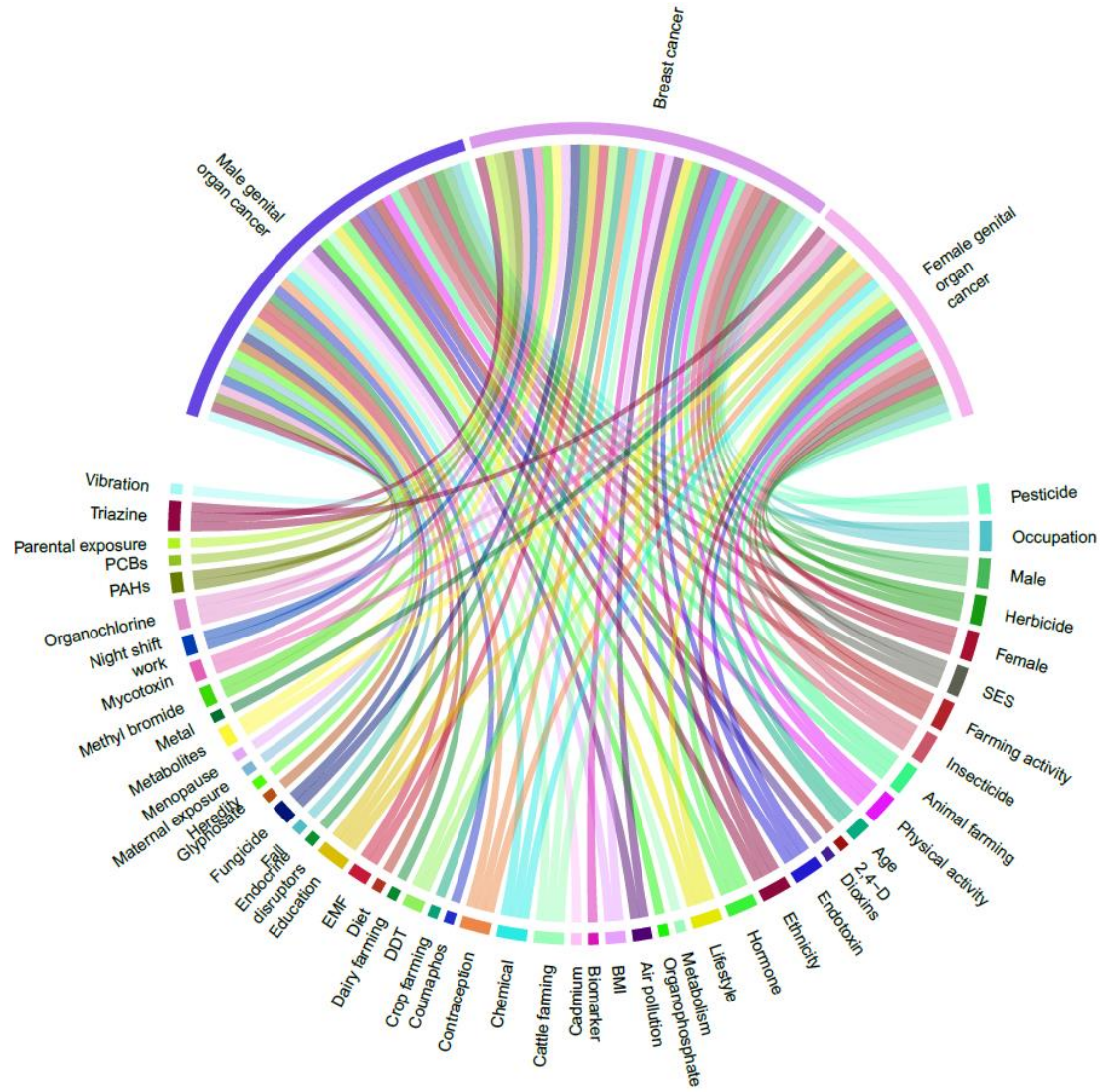


368 **Figure S22.** Keyword co-occurrence network between potential risk factor and autoimmune disease keywords.
 369 BMI: body mass index, BP: blood pressure, IBD: inflammatory bowel disease, SES: socio-economic status, VOC: volatile organic compound. Larger the node size, higher the
 370 number of co-occurrence. Red nodes refer to health events while blue nodes refer to potential risk factors.

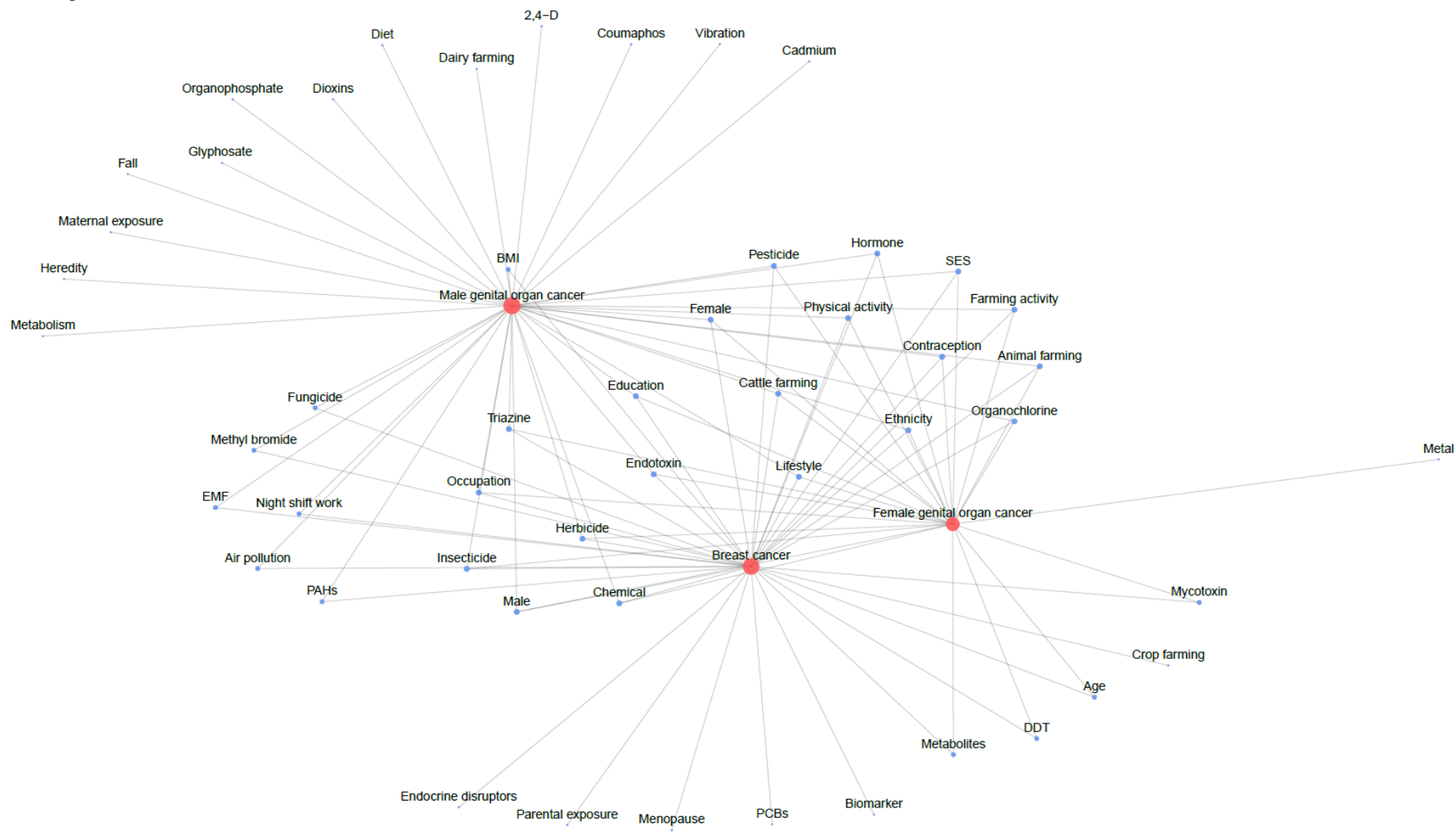


371
 372

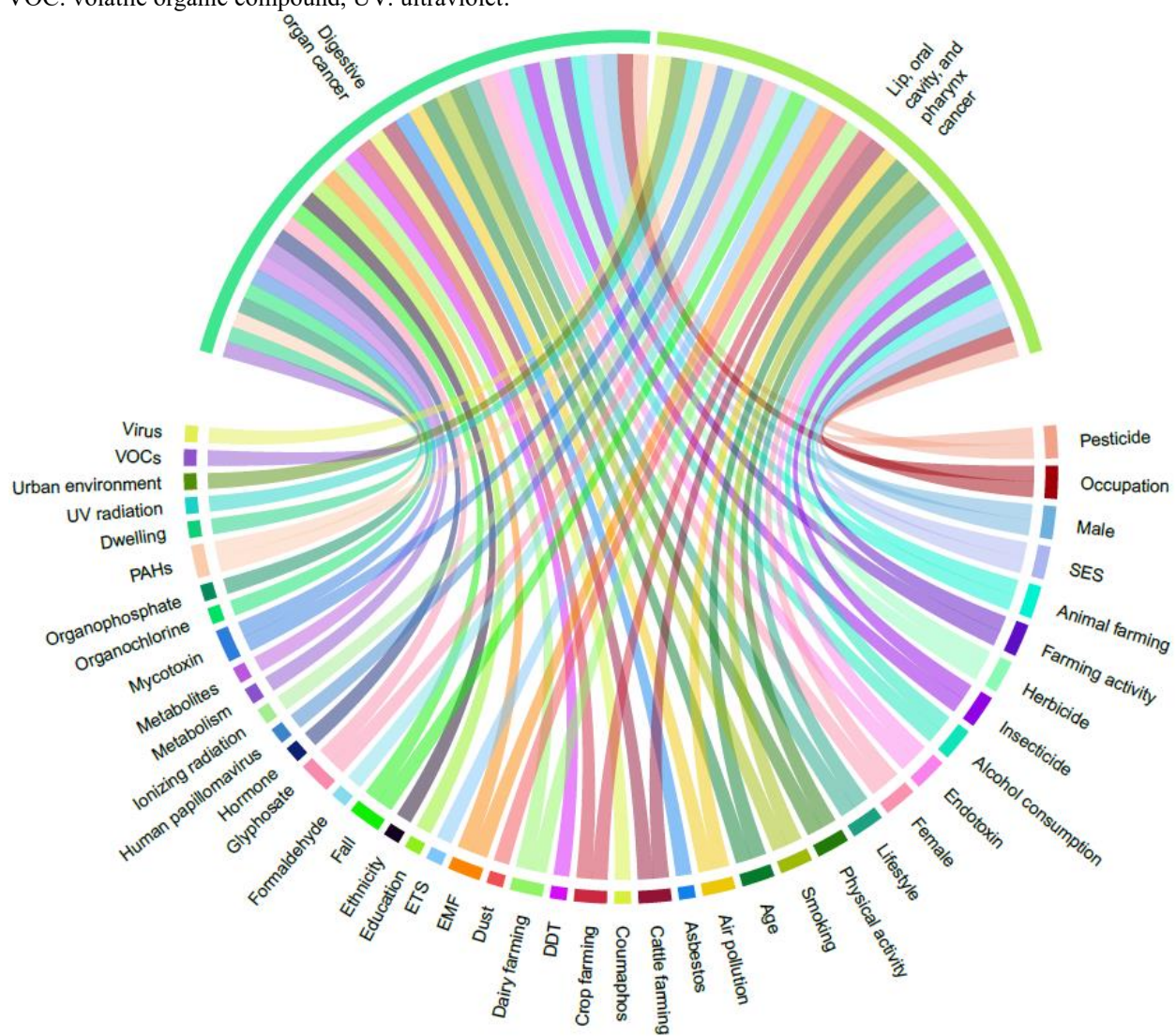
373 **Figure S23.** Chord diagram of keyword co-occurrence between potential risk factor and breast and genital organ cancer keywords.
 374 2,4-D: 2,4-dichlorophenoxyacetic acid, BMI: body mass index, DDT: dichlorodiphenyltrichloroethane, EMF: electromagnetic field, PAH: polycyclic aromatic hydrocarbon,
 375 PCB: polychlorinated biphenyls, SES: socio-economic status.



377 **Figure S24.** Keyword co-occurrence network between potential risk factor and breast and genital organ cancer keywords.
 378 2,4-D: 2,4-dichlorophenoxyacetic acid, BMI: body mass index, DDT: dichlorodiphenyltrichloroethane, EMF: electromagnetic field, PAH: polycyclic aromatic hydrocarbon,
 379 PCB: polychlorinated biphenyls, SES: socio-economic status. Larger the node size, higher the number of co-occurrence. Red nodes refer to health events while blue nodes
 380 refer to potential risk factors.

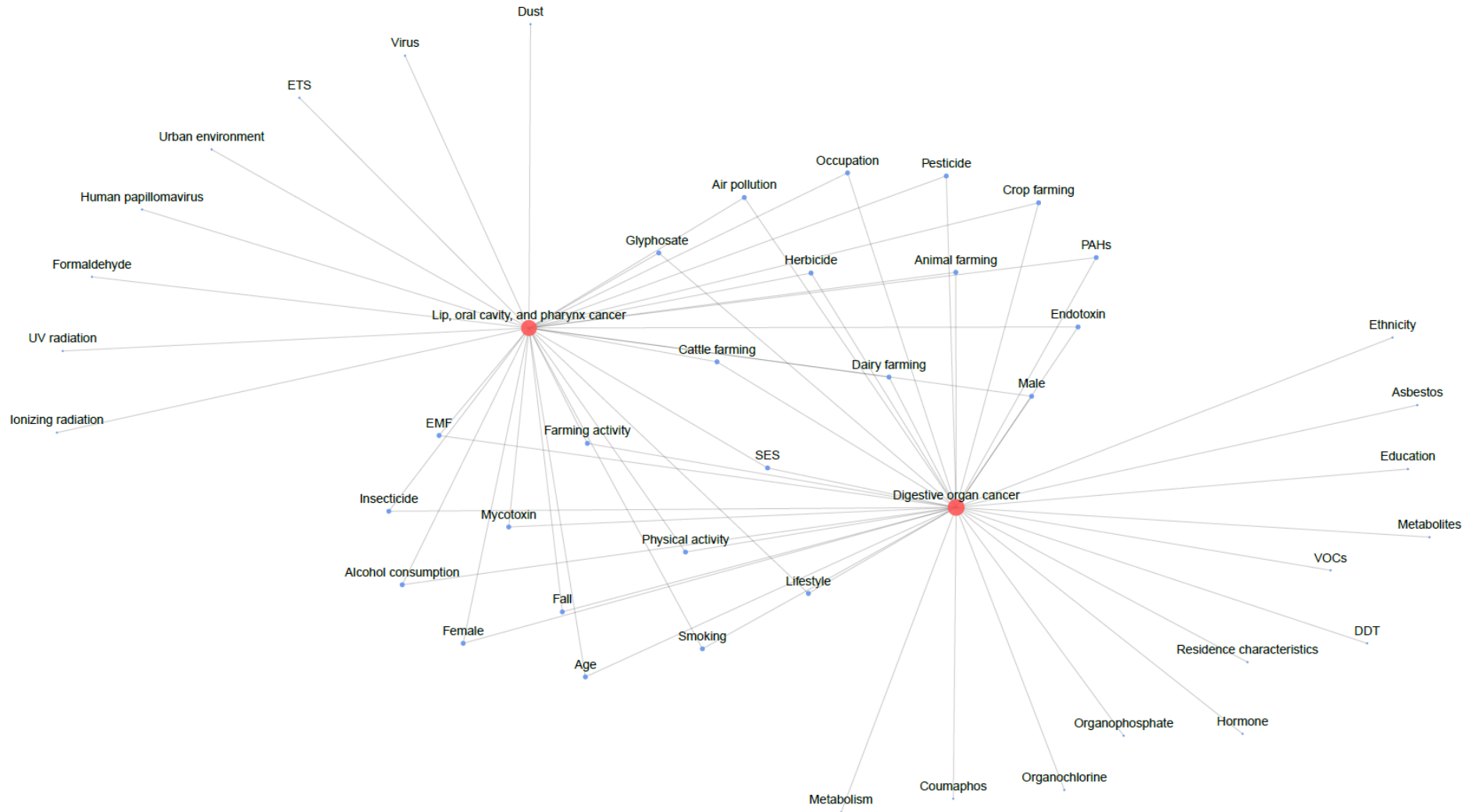


382 **Figure S25.** Chord diagram of keyword co-occurrence between potential risk factor and digestive organ, lip, oral cavity and pharynx cancer keywords.
 383 BMI: body mass index, DDT: dichlorodiphenyltrichloroethane, EMF: electromagnetic field, ETS: environmental tobacco smoking, PAH: polycyclic aromatic hydrocarbon,
 384 SES: socio-economic status, VOC: volatile organic compound, UV: ultraviolet.

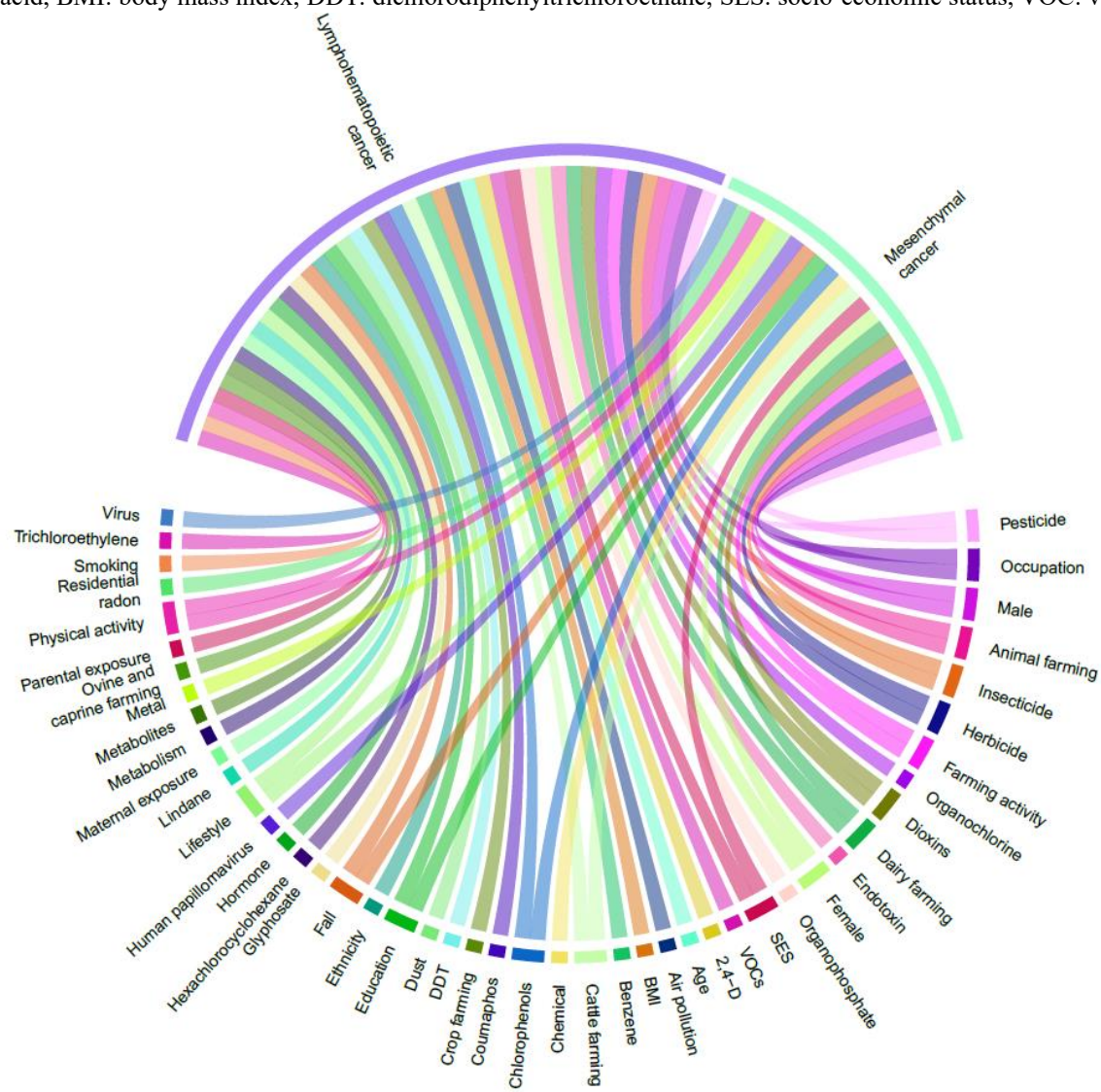


385

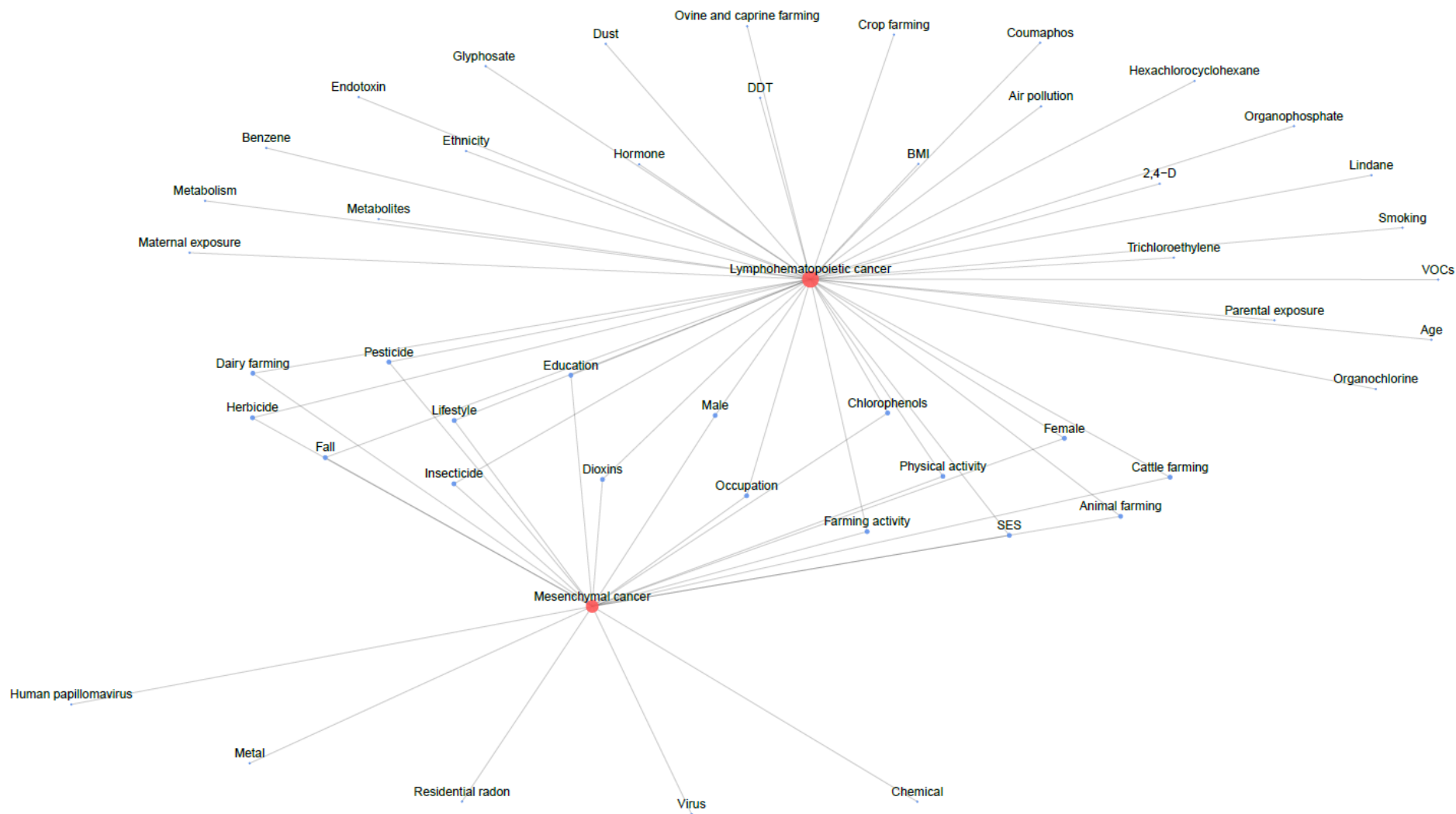
386 **Figure S26.** Keyword co-occurrence network between potential risk factor and digestive organ, lip, oral cavity and pharynx cancer keywords.
 387 BMI: body mass index, DDT: dichlorodiphenyltrichloroethane, EMF: electromagnetic field, ETS: environmental tobacco smoking, PAH: polycyclic aromatic hydrocarbon,
 388 SES: socio-economic status, VOC: volatile organic compound, UV: ultraviolet. Larger the node size, higher the number of co-occurrence. Red nodes refer to health events
 389 while blue nodes refer to potential risk factors.



391 **Figure S27.** Chord diagram of keyword co-occurrence between potential risk factor and lymphohematopoietic and mesenchymal cancer keywords.
 392 2,4-D: 2,4-dichlorophenoxyacetic acid, BMI: body mass index, DDT: dichlorodiphenyltrichloroethane, SES: socio-economic status, VOC: volatile organic compound.

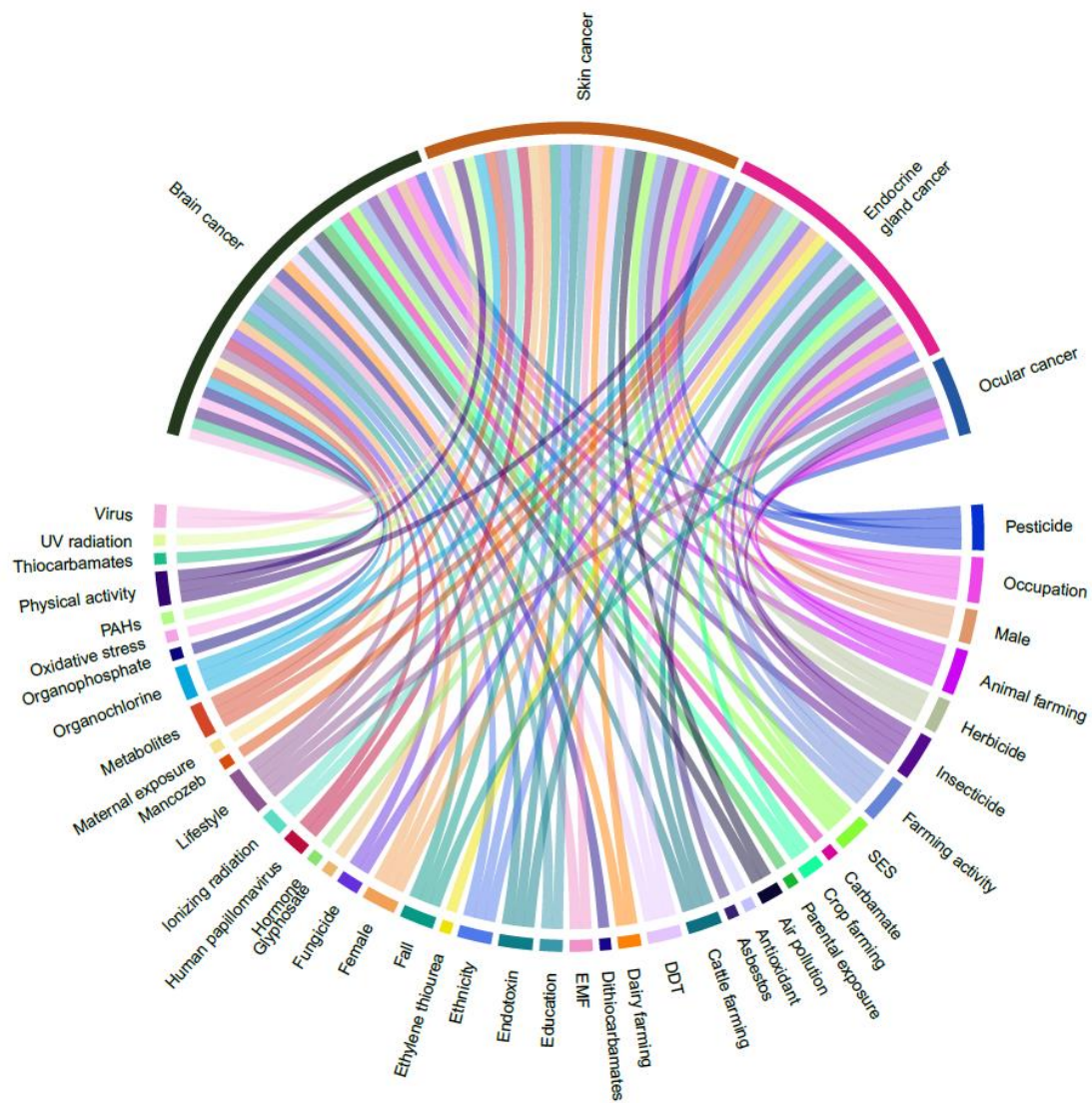


394 **Figure S28.** Keyword co-occurrence network between potential risk factor and lymphohematopoietic and mesenchymal cancer keywords.
 395 2,4-D: 2,4-dichlorophenoxyacetic acid, BMI: body mass index, DDT: dichlorodiphenyltrichloroethane, SES: socio-economic status, VOC: volatile organic compound. Larger
 396 the node size, higher the number of co-occurrence. Red nodes refer to health events while blue nodes refer to potential risk factors.



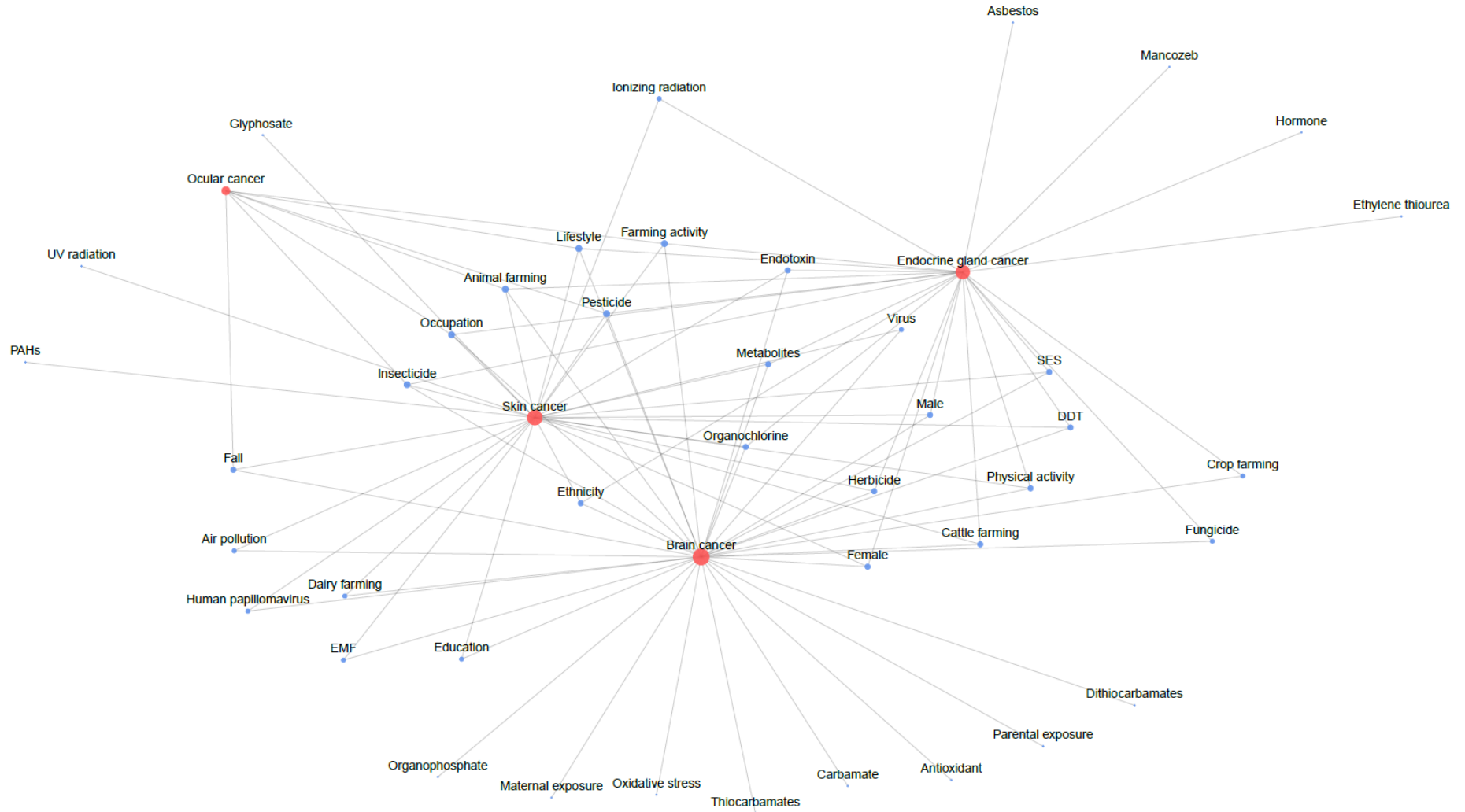
397

398 **Figure S29.** Chord diagram of keyword co-occurrence between potential risk factor and brain, skin, ocular and endocrine gland cancer keywords.
 399 DDT: dichlorodiphenyltrichloroethane, EMF: electromagnetic field, PAH: polycyclic aromatic hydrocarbon, SES: socio-economic status, UV: ultraviolet.



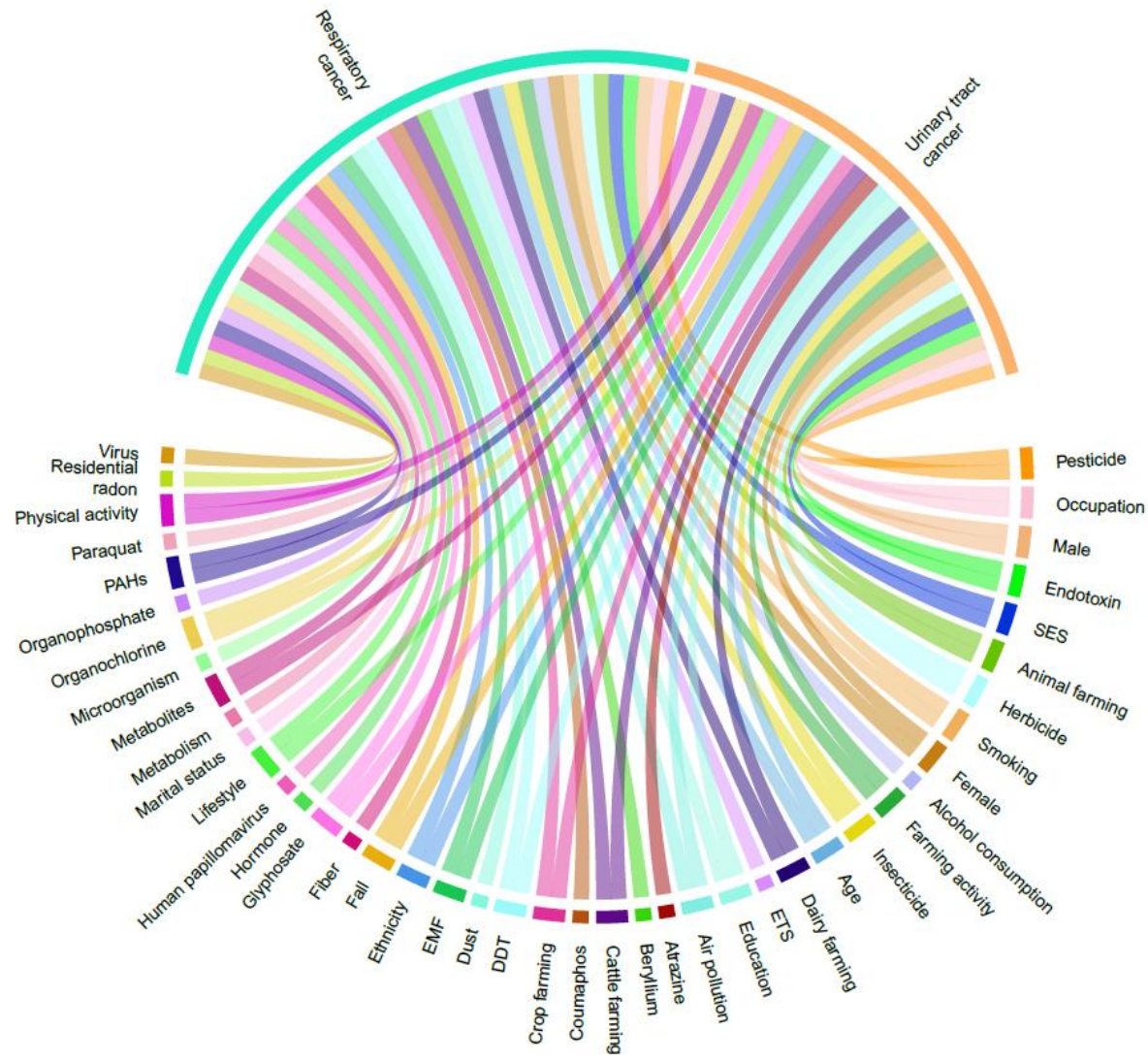
400
 401

402 **Figure S30.** Keyword co-occurrence network between potential risk factor and brain, skin, ocular and endocrine gland cancer keywords.
 403 DDT: dichlorodiphenyltrichloroethane, EMF: electromagnetic field, PAH: polycyclic aromatic hydrocarbon, SES: socio-economic status, UV: ultraviolet. Red nodes refer to
 404 health events while blue nodes refer to potential risk factors.

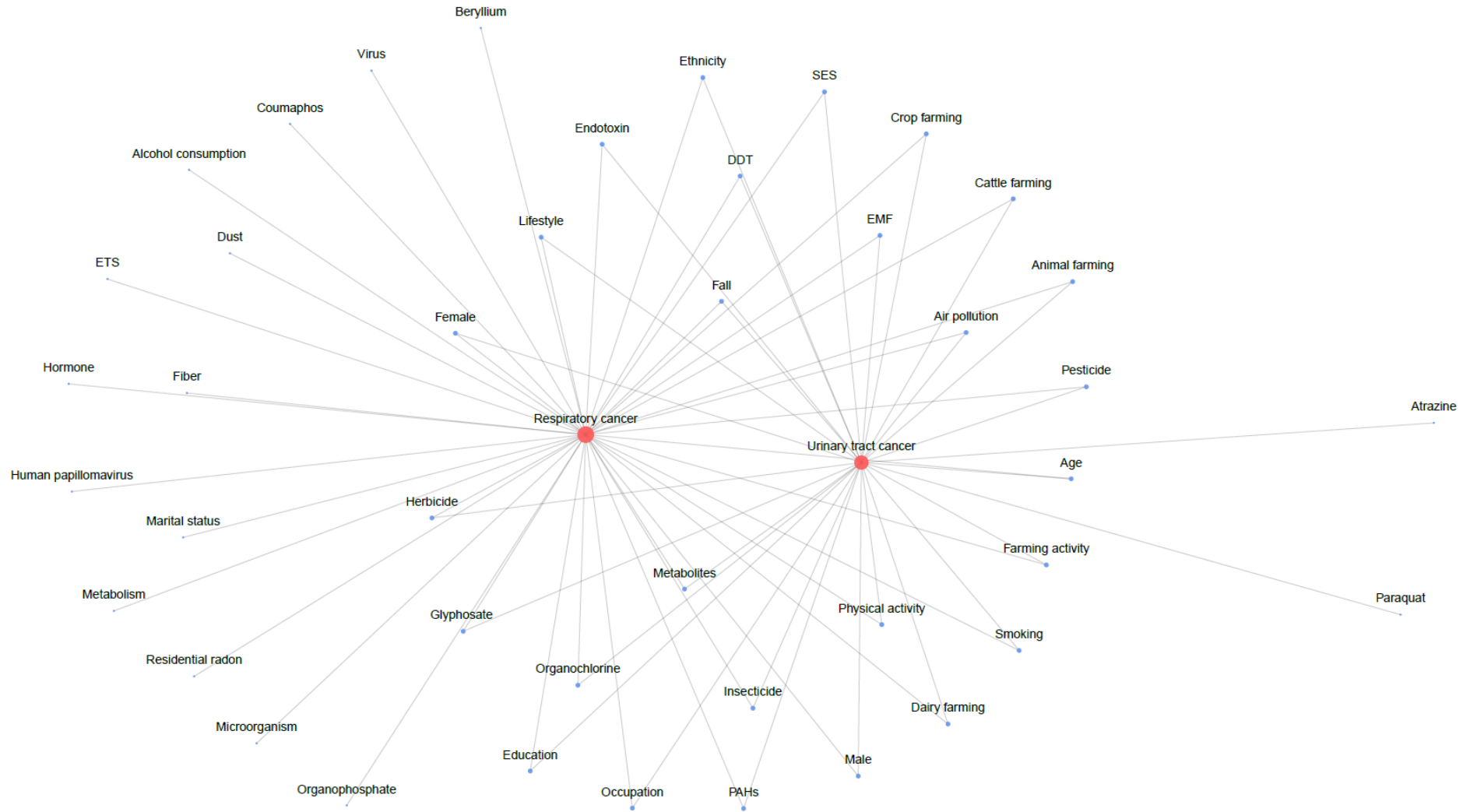


405
406

407 **Figure S31.** Chord diagram of keyword co-occurrence between potential risk factor and respiratory and urinary tract cancer keywords.
 408 DDT: dichlorodiphenyltrichloroethane, EMF: electromagnetic field, ETS: environmental tobacco smoking, PAH: polycyclic aromatic hydrocarbon, SES: socio-economic
 409 status.

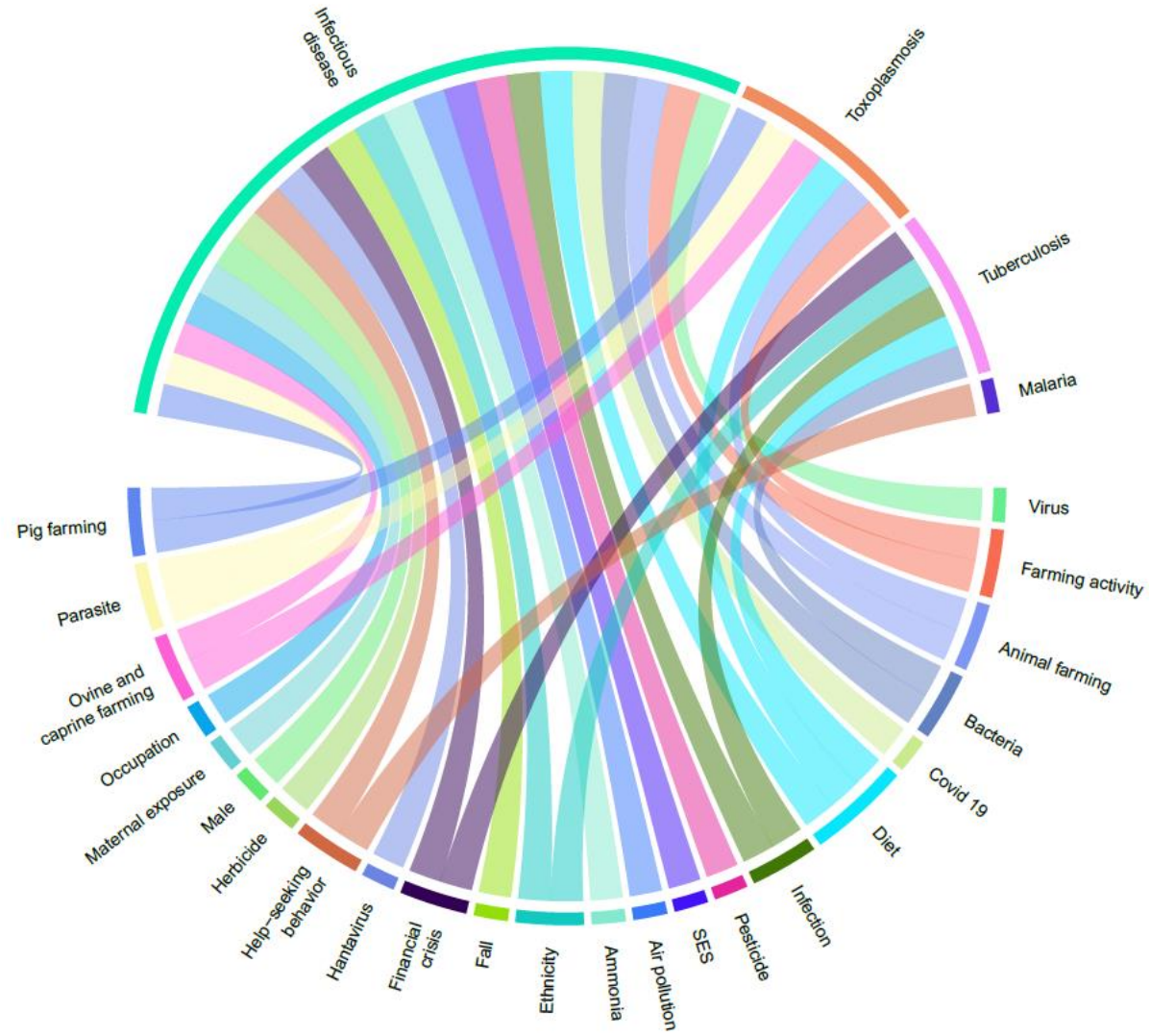


411 **Figure S32.** Keyword co-occurrence network between potential risk factor and respiratory and urinary tract cancer keywords.
412 DDT: dichlorodiphenyltrichloroethane, EMF: electromagnetic field, ETS: environmental tobacco smoking, PAH: polycyclic aromatic hydrocarbon, SES: socio-economic
413 status. Larger the node size, higher the number of co-occurrence. Red nodes refer to health events while blue nodes refer to potential risk factors.



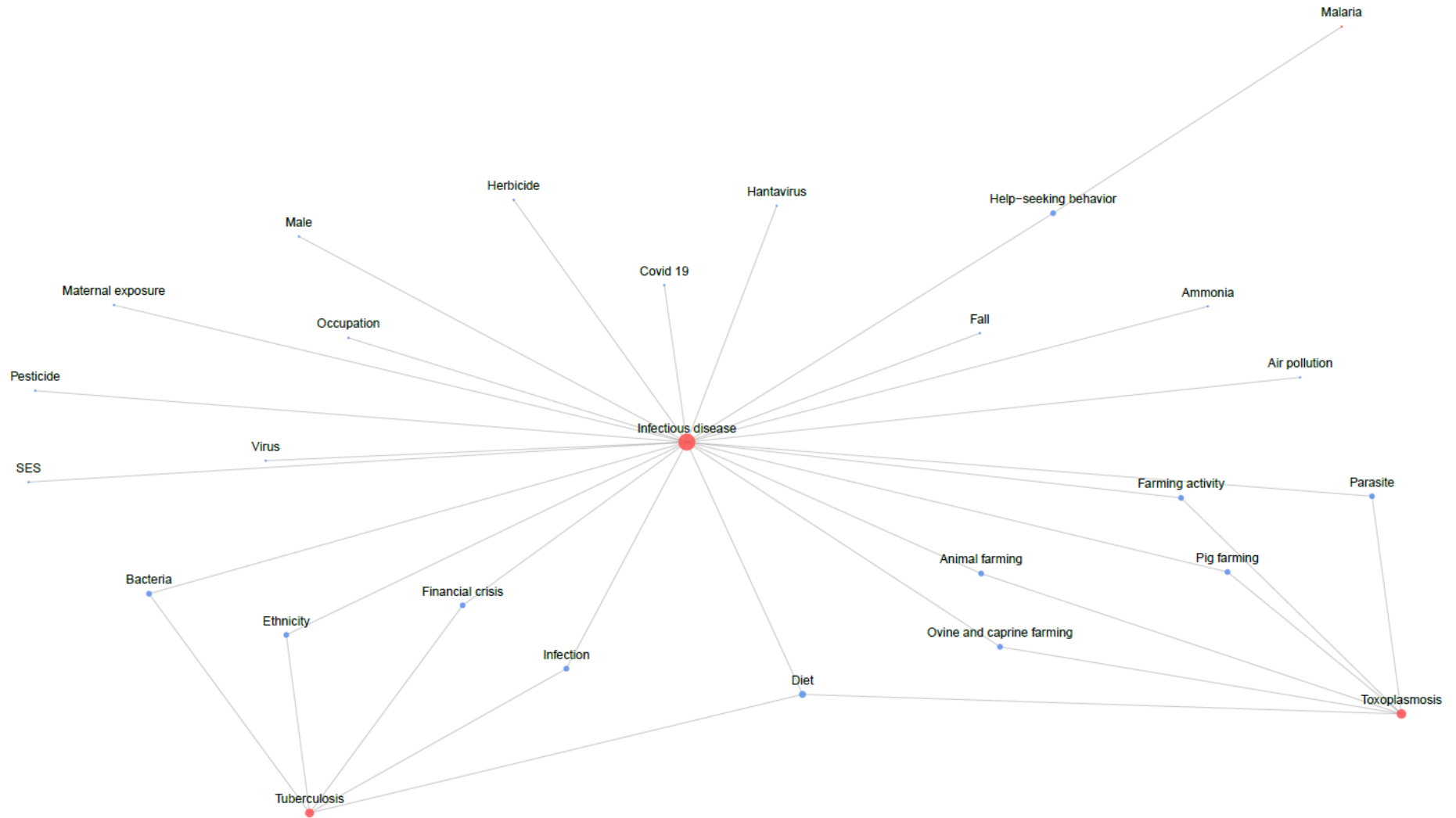
414
415

416 **Figure S33.** Chord diagram of keyword co-occurrence between potential risk factor and infectious disease keywords.
417 SES: socio-economic status.



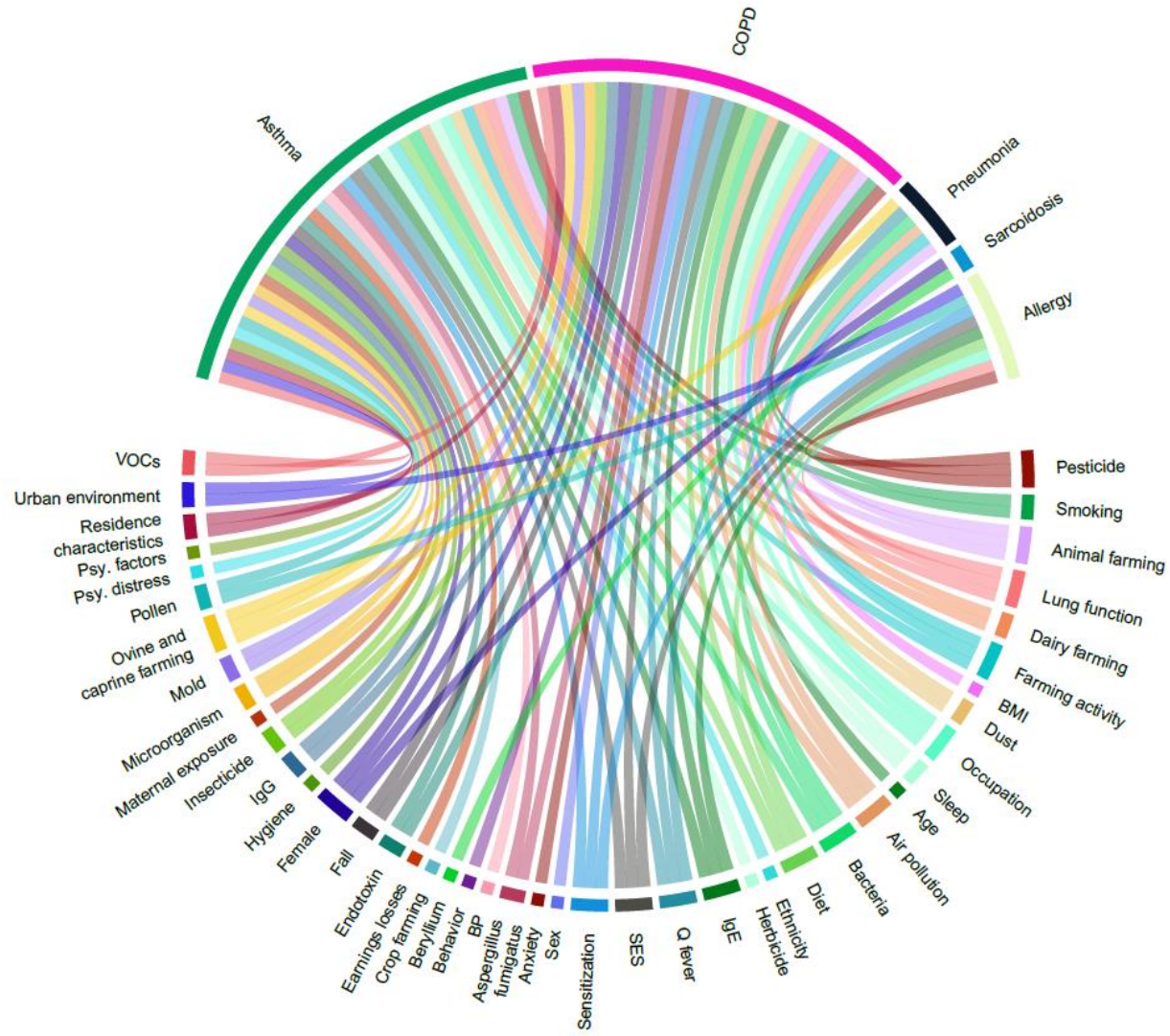
418
419

420 **Figure S34.** Keyword co-occurrence network between potential risk factor and infectious disease keywords.
421 SES: socio-economic status. Larger the node size, higher the number of co-occurrence. Red nodes refer to health events while blue nodes refer to potential risk factors.

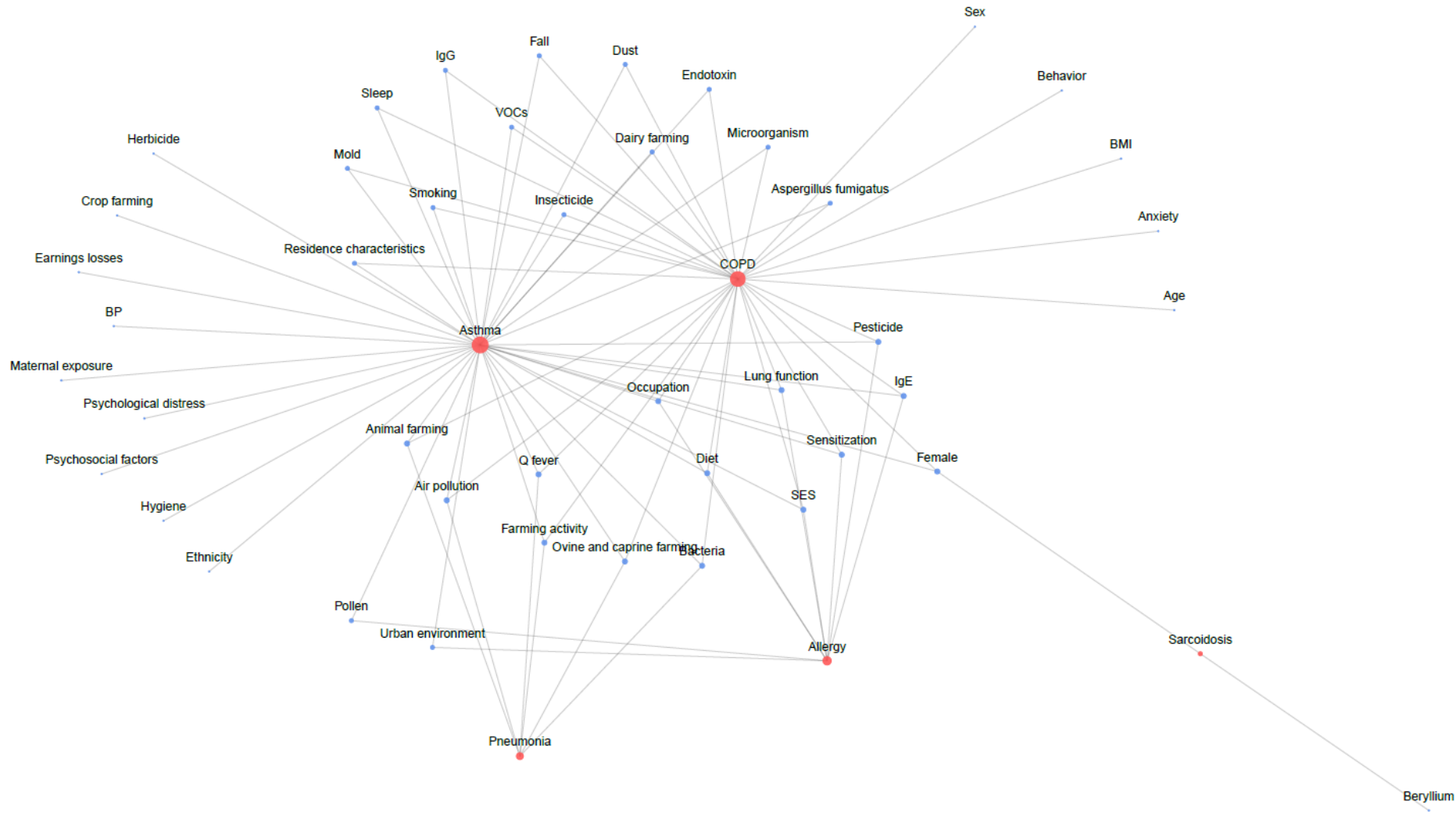


422
423

424 **Figure S35.** Chord diagram of keyword co-occurrence between potential risk factor and respiratory disorder keywords.
 425 BMI: body mass index, BP: blood pressure, COPD: chronic obstructive pulmonary disease, IgE: immunoglobulin E, IgG: immunoglobulin G, psy: psychosocial, SES: socio-
 426 economic status.

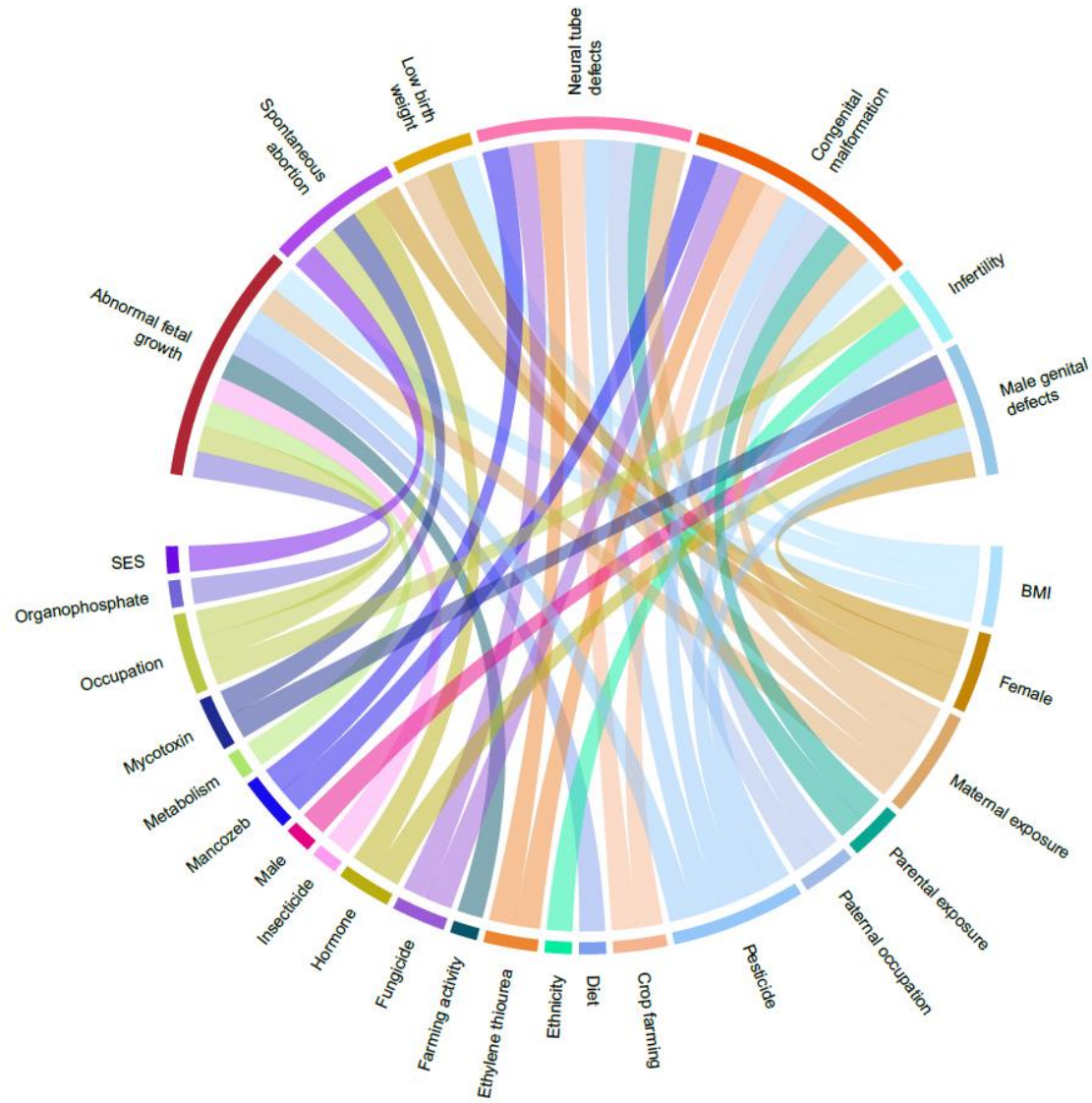


428 **Figure S36.** Keyword co-occurrence network between potential risk factor and respiratory disorder keywords.
429 BMI: body mass index, BP: blood pressure, COPD: chronic obstructive pulmonary disease, IgE: immunoglobulin E, IgG: immunoglobulin G, psy: psychosocial, SES: socio-
430 economic status. Larger the node size, higher the number of co-occurrence. Red nodes refer to health events while blue nodes refer to potential risk factors.

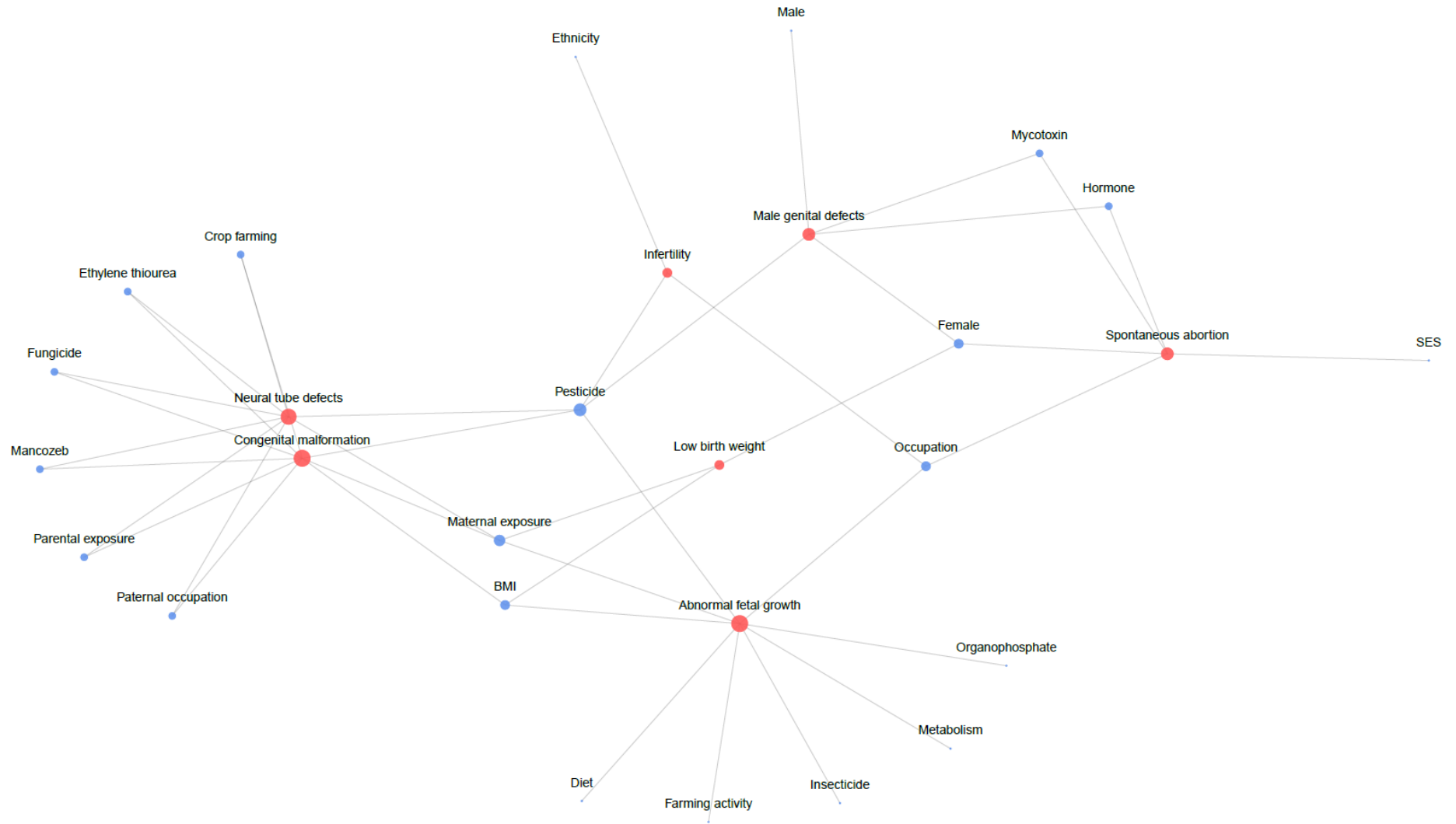


431
432

433 **Figure S37.** Chord diagram of keyword co-occurrence between potential risk factor and reproductive disorder keywords.
 434 BMI: body mass index, SES: socio-economic status.

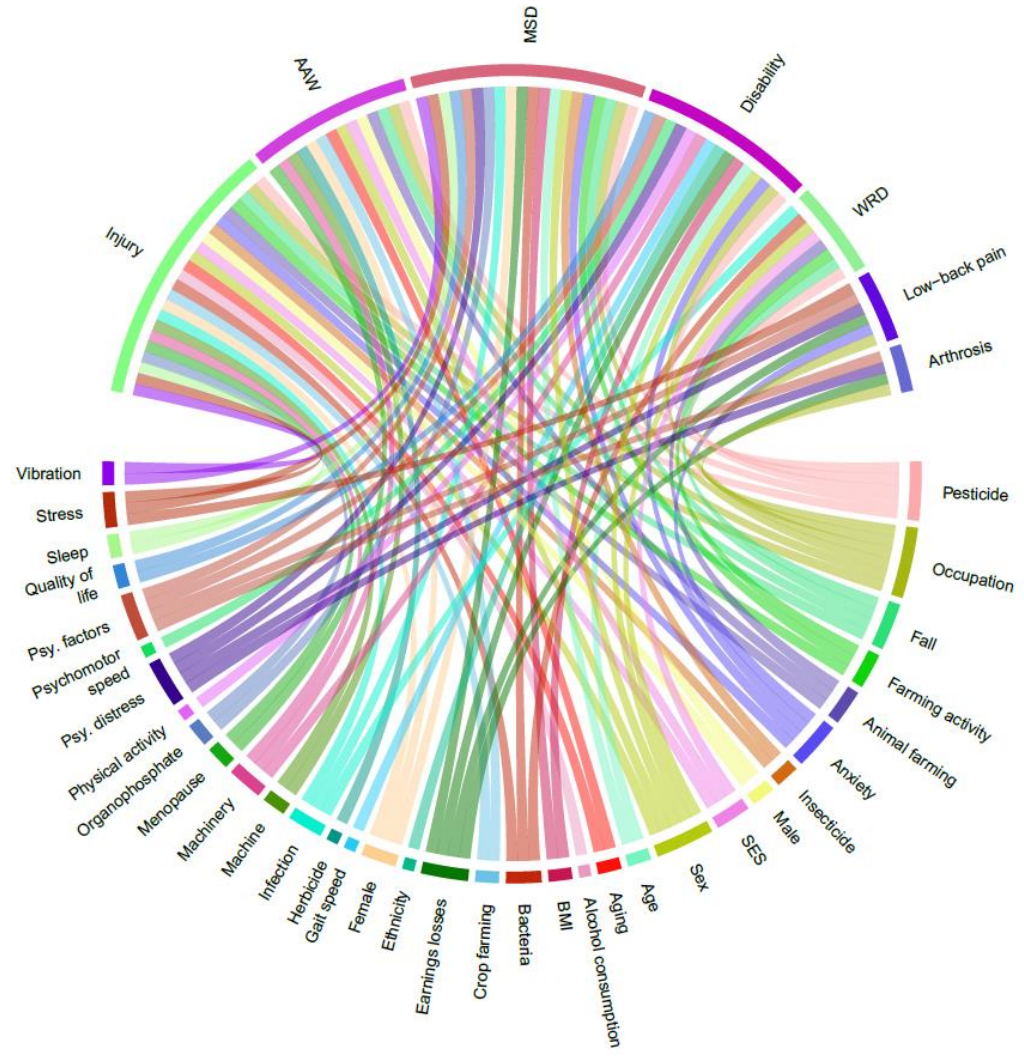


436 **Figure S38.** Keyword co-occurrence network between potential risk factor and reproductive disorder keywords.
 437 BMI: body mass index, SES: socio-economic status. Larger the node size, higher the number of co-occurrence. Red nodes refer to health events while blue nodes refer to
 438 potential risk factors.



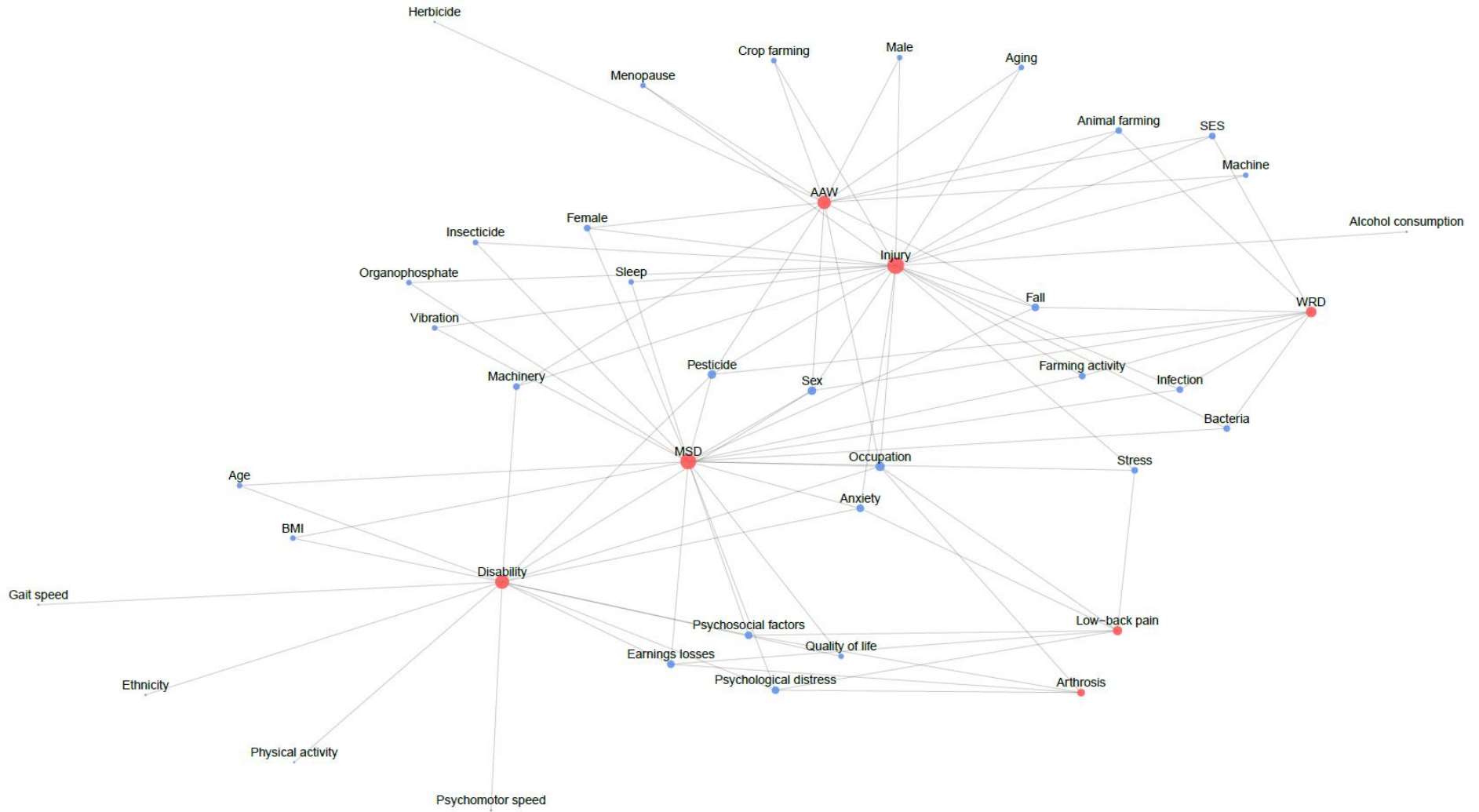
439
 440

441 **Figure S39.** Chord diagram of keyword co-occurrence between potential risk factor and injury and work-related disease keywords.
 442 AAW: workplace accident, BMI: body mass index, MSD: musculoskeletal disorder, psy: psychological, SES: socio-economic status, WRD: work-related disease.



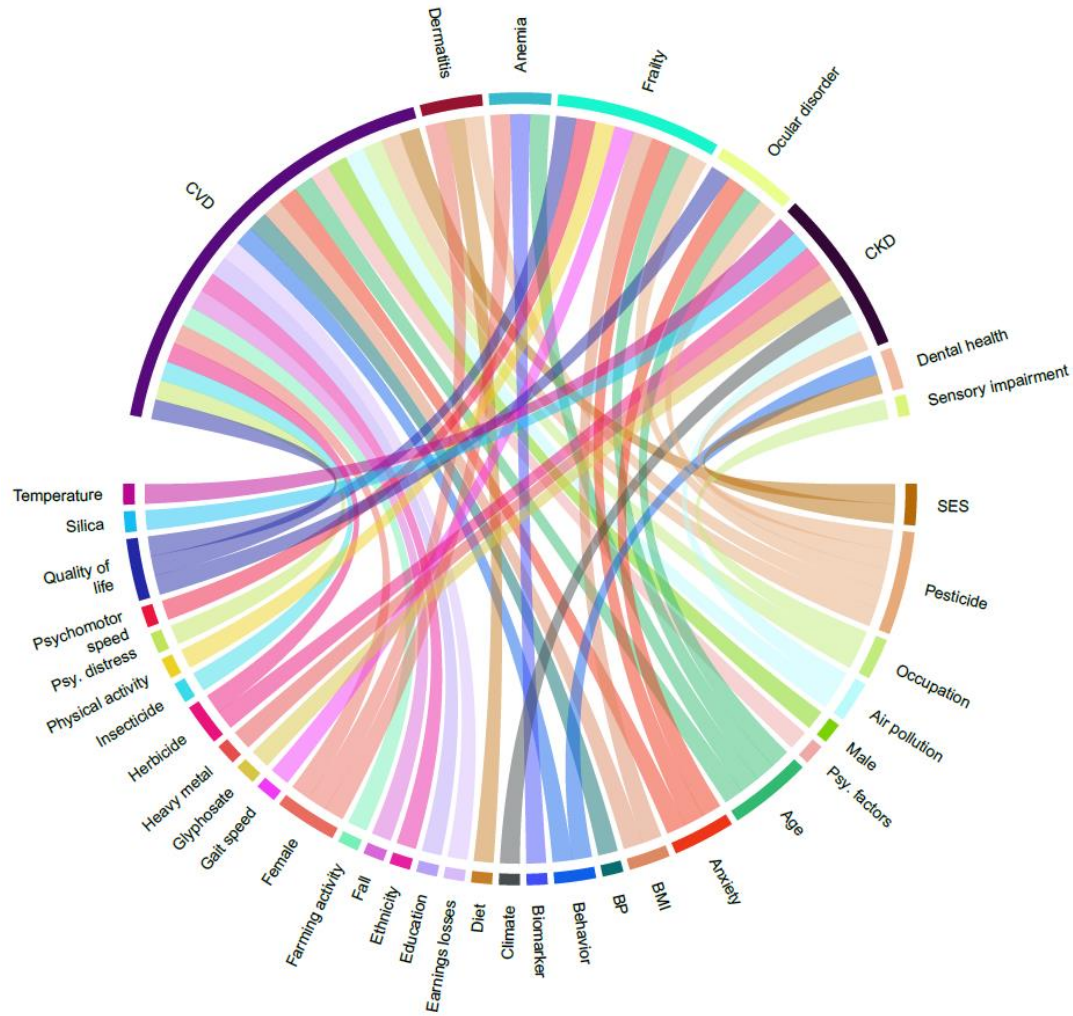
443
 444

445 **Figure S40.** Keyword co-occurrence network between potential risk factor and injury and work-related disease keywords.
 446 AAW: workplace accident, BMI: body mass index, MSD: musculoskeletal disorder, psy: psychological, SES: socio-economic status, WRD: work-related disease. Larger the
 447 node size, higher the number of co-occurrence. Red nodes refer to health events while blue nodes refer to potential risk factors.



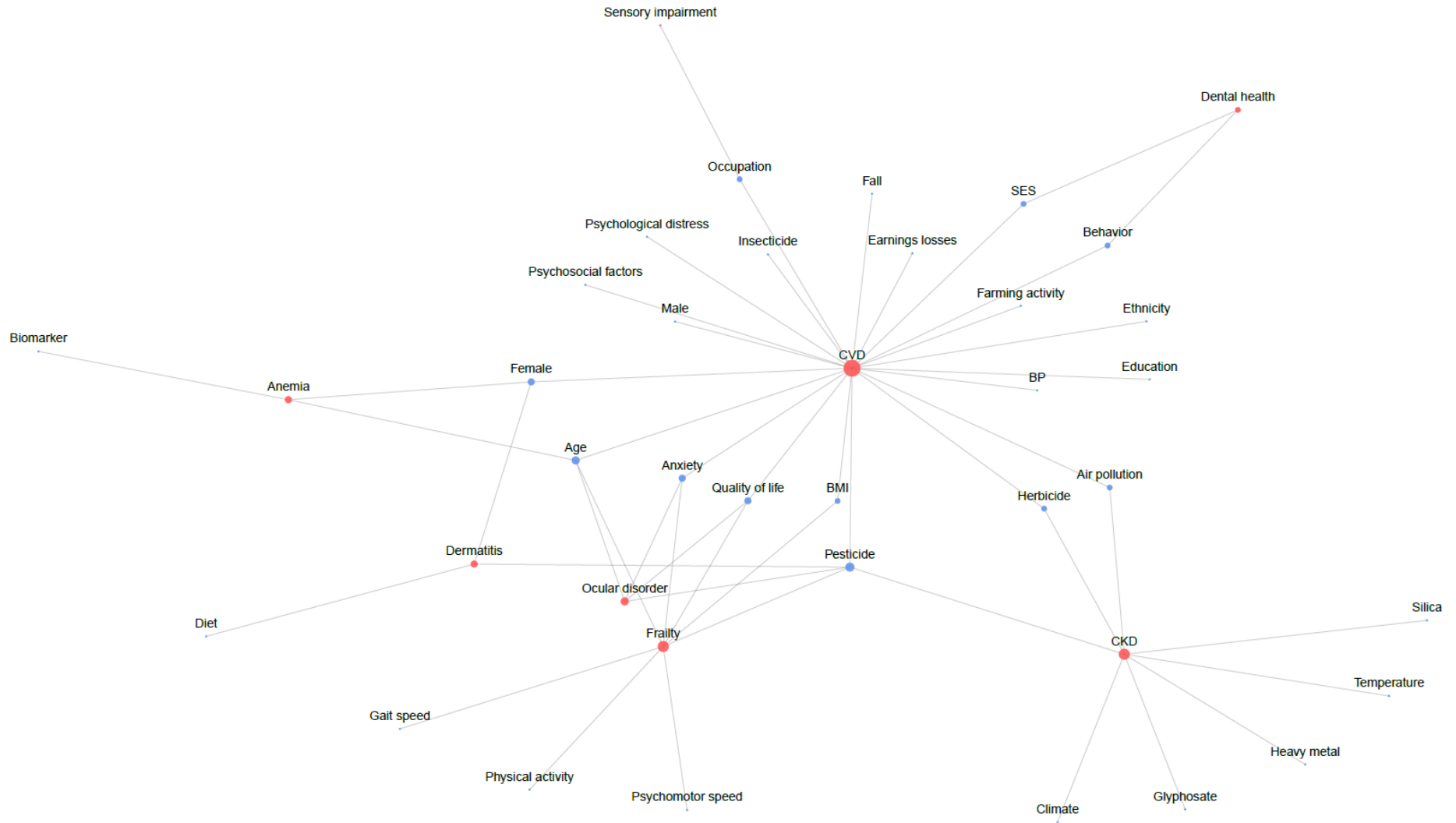
448
 449

450 **Figure S41.** Chord diagram of keyword co-occurrence between potential risk factor and other health event keywords.
 451 BMI: body mass index, BP: blood pressure, CKD: chronic kidney disease, CVD: cardiovascular disease, psy: psychological, SES: socio-economic status.



452
 453

454 **Figure S42.** Keyword co-occurrence network between potential risk factor and other health event keywords.
 455 BMI: body mass index, BP: blood pressure, CKD: chronic kidney disease, CVD: cardiovascular disease, psy: psychological, SES: socio-economic status. Larger the node
 456 size, higher the number of co-occurrence. Red nodes refer to health events while blue nodes refer to potential risk factors.



457
458

463 **Abbreviations**

- 464
465 90%CI: 90% confidence interval
466 AAW: workplace accident
467 AGEIS: Autonomy, Gerontology, E-health, Imaging & Society (Autonomie, G erontologie, E-sant e, Imagerie et
468 Soci et e in French)
469 AGR: annual growth rate
470 AGRICAN: Agriculture and Cancer
471 AHD: administrative health database
472 AHS: Agricultural Health Study
473 AI: artificial intelligence
474 AMI: Aging Multidisciplinary Investigation
475 ANR: French National Research Agency (Agence Nationale de la Recherche in French)
476 AS: association strength
477 BIBLIO: preliminary guideline for reporting bibliometric reviews of the biomedical literature
478 BMI: body mass index
479 CHU: university hospital (centre hospitalier universitaire in French)
480 COPD: chronic obstructive pulmonary disease
481 CNAP: Cancer in the Norwegian Agricultural population cohort
482 CVD: cardiovascular disorder
483 DDT: dichlorodiphenyltrichloroethane
484 EHR: electronic health record
485 EMR: electronic medical record
486 FAIR: findable, accessible, interoperable and reusable
487 FERMA: risk factors of the rural environment and the allergic and respiratory disease
488 FF: fractional frequency
489 FINJEM: Finnish information system on occupational exposure
490 GDP: gross domestic product
491 IARC: International Agency for Research on Cancer
492 IBD: inflammatory bowel disease
493 ICMJE: International Committee of Medical Journal Editors
494 INCa: French national cancer institute (Institut National du Cancer in French)
495 INSERM: French National Institute of Health and Medical Research (Institut national de la sant e et de la recherche
496 m edicale in French)
497 IQR: interquartile range
498 JEM: job-exposure matrix
499 mHealth: mobile health
500 MIAI: Multidisciplinary Institute in Artificial Intelligence
501 ML: machine learning
502 MSA: National Health Insurance Fund for Agricultural Workers and Farmers (Mutualit e Sociale Agricole in
503 French)
504 MSD: musculoskeletal disorder
505 NAWS: National Agricultural Workers Survey
506 NCI: National Cancer Institute
507 NIH: National Institutes of Health
508 NOCCA: Nordic Occupational Cancer Study
509 ORCID: Open Researcher and Contributor ID
510 PFAS: per- and polyfluoroalkyl substances
511 PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping
512 Reviews and evidence maps
513 Q1: first quartile
514 R&D: research and development
515 SES: socio-economic status
516 TEDI: toxicological and exposure database inventory
517 TRACTOR: Tracking and monitoring occupational risks in agriculture
518 UFW: United Farm Workers of America
519 UGA: Universit e Grenoble Alpes
520 UK: United Kingdom
521 US: United States
522 UV: ultraviolet

523 **References**

524

- 525 1. Bienert IR, Oliveira RC, Andrade PB, Caramori CA. Bibliometric indexes, databases and impact factors
526 in cardiology. *Rev Bras Cir Cardiovasc* 2015;30(2):254–259. doi:10.5935/1678-9741.20150019.
- 527 2. Bihari A, Tripathi S, Deepak A. A review on h-index and its alternative indices. *J Inf Sci* 2023;49(3):624–
528 665 (2023). doi:10.1177/0165551521101447.
- 529 3. Shao Y, Chien TW, Jang FL. The use of radar plots with the Yk-index to identify which authors
530 contributed the most to the journal of Medicine in 2020 and 2021: A bibliometric analysis. *Medicine*
531 (Baltimore) 2022;101(45):e31033. doi:10.1097%2FMD.00000000000031033.
- 532 4. Kumar S, Kumar S. Trends of collaborative research in journals of oilseeds research (India), 1993-2004.
533 *Indian J Agric Library Inf Services* 2008;24:80–90.
- 534 5. Wu T, Duan Y, Zhang T, Tian W, Liu H, Deng Y. Research Trends in the Application of Artificial
535 Intelligence in Oncology: A Bibliometric and Network Visualization Study. *Front Biosci (Landmark Ed)*.
536 2022;27(9):254 doi:10.31083/j.fbl2709254.
- 537 6. Pukkala E, Martinsen JI, Lyng E, et al. Occupation and cancer - follow-up of 15 million people in five
538 Nordic countries. *Acta Oncol*. 2009;48(5):646-790 doi:10.1080/02841860902913546.
- 539 7. Alavanja MC, Samanic C, Dosemeci M, et al. Use of agricultural pesticides and prostate cancer risk in
540 the Agricultural Health Study cohort. *Am J Epidemiol* 2003;157(9):800-814. doi:10.1093/aje/kwg040.
- 541 8. Elbaz A, Clavel J, Rathouz PJ, et al. Professional exposure to pesticides and Parkinson disease. *Ann*
542 *Neurol* 2009;66(4):494-504. doi:10.1002/ana.21717.
- 543 9. Andersen A, Barlow L, Engeland A, Kjaerheim K, Lyng E, Pukkala E. Work-related cancer in the
544 Nordic countries. *Scand J Work Environ Health* 1999;25 Suppl 2:1-116. PMID: 10507118.
- 545 10. Karjalainen A, Kurppa K, Martikainen R, Klaukka T, Karjalainen J. Work is related to a substantial
546 portion of adult-onset asthma incidence in the Finnish population. *Am J Respir Crit Care Med*
547 2001;164(4):565-568. doi:10.1164/ajrccm.164.4.2012146.
- 548 11. Vingård E, Alfredsson L, Goldie I, Hogstedt C. Occupation and osteoarthritis of the hip and knee: a
549 register-based cohort study. *Int J Epidemiol* 1991;20(4):1025-1031. doi:10.1093/ije/20.4.1025.
- 550 12. Alavanja MC, Dosemeci M, Samanic C, et al. Pesticides and lung cancer risk in the agricultural health
551 study cohort. *Am J Epidemiol* 2004;160(9):876-885. doi:10.1093/aje/kwh290.
- 552 13. Andreotti G, Koutros S, Hofmann JN, et al. Glyphosate Use and Cancer Incidence in the Agricultural
553 Health Study. *J Natl Cancer Inst* 2018;110(5):509-516. doi:10.1093/jnci/djx233.
- 554 14. Wigle DT, Semenciw RM, Wilkins K, et al. Mortality study of Canadian male farm operators: non-
555 Hodgkin's lymphoma mortality and agricultural practices in Saskatchewan. *J Natl Cancer Inst*
556 1990;82(7):575-582. doi:10.1093/jnci/82.7.575.
- 557 15. Engel LS, Hill DA, Hoppin JA, et al. Pesticide use and breast cancer risk among farmers' wives in the
558 agricultural health study. *Am J Epidemiol* 2005;161(2):121-135. doi:10.1093/aje/kwi022.
- 559 16. Hemminki K, Saloniemi I, Salonen T, Partanen T, Vainio H. Childhood cancer and parental occupation
560 in Finland. *J Epidemiol Community Health* 1981;35(1):11-15. doi:10.1136/jech.35.1.11.
- 561 17. Weichenthal S, Villeneuve PJ, Burnett RT, et al. Long-term exposure to fine particulate matter:
562 association with nonaccidental and cardiovascular mortality in the agricultural health study cohort.
563 *Environ Health Perspect* 2014;122(6):609-615. doi:10.1289/ehp.1307277.

- 564 18. Levecque C, Elbaz A, Clavel J, et al. Association between Parkinson's disease and polymorphisms in the
565 nNOS and iNOS genes in a community-based case-control study. *Hum Mol Genet* 2003;12(1):79-86.
566 doi:10.1093/hmg/ddg009.
- 567 19. Morrison H, Savitz D, Semenciw R, et al. Farming and prostate cancer mortality. *Am J Epidemiol*
568 1993;137(3):270-280. doi:10.1093/oxfordjournals.aje.a116674.
- 569 20. Kristensen P, Andersen A, Irgens LM, Laake P, Bye AS. Incidence and risk factors of cancer among men
570 and women in Norwegian agriculture. *Scand J Work Environ Health* 1996;22(1):14-26.
571 doi:10.5271/sjweh.104.
- 572 21. Orton SM, Wald L, Confavreux C, et al. Association of UV radiation with multiple sclerosis prevalence
573 and sex ratio in France. *Neurology* 2011;76(5):425-431. doi:10.1212/WNL.0b013e31820a0a9f.
- 574 22. Stokes L, Stark A, Marshall E, Narang A. Neurotoxicity among pesticide applicators exposed to
575 organophosphates. *Occup Environ Med* 1995;52(10):648-653. doi:10.1136/oem.52.10.648.
- 576 23. Vukusic S, Van Bockstael V, Gosselin S, Confavreux C. Regional variations in the prevalence of multiple
577 sclerosis in French farmers. *J Neurol Neurosurg Psychiatry* 2007;78(7):707-709.
578 doi:10.1136/jnnp.2006.101196.
- 579 24. Mills PK, Yang R. Prostate cancer risk in California farm workers. *J Occup Environ Med* 2003;45(3):249-
580 258. doi:10.1097/01.jom.0000058339.05741.0c.
- 581 25. Eriksson M, Karlsson M. Occupational and other environmental factors and multiple myeloma: a
582 population based case-control study. *Br J Ind Med* 1992;49(2):95-103. doi:10.1136/oem.49.2.95.
- 583 26. Wiklund K, Dich J, Holm LE, Eklund G. Risk of cancer in pesticide applicators in Swedish agriculture.
584 *Br J Ind Med* 1989;46(11):809-814. doi:10.1136/oem.46.11.809.
- 585 27. Pouchieu C, Piel C, Carles C, et al. Pesticide use in agriculture and Parkinson's disease in the AGRICAN
586 cohort study. *Int J Epidemiol* 2018;47(1):299-310. doi:10.1093/ije/dyx225.
- 587 28. Karjalainen A, Martikainen R, Klaukka T, Saarinen K, Uitti J. Risk of asthma among Finnish patients
588 with occupational rhinitis. *Chest* 2003;123(1):283-288. doi:10.1378/chest.123.1.283.
- 589 29. Leon ME, Schinasi LH, Lebailly P, et al. Pesticide use and risk of non-Hodgkin lymphoid malignancies
590 in agricultural cohorts from France, Norway and the USA: a pooled analysis from the AGRICOH
591 consortium. *Int J Epidemiol* 2019;48(5):1519-1535. doi:10.1093/ije/dyz017.
- 592 30. Lee WJ, Sandler DP, Blair A, Samanic C, Cross AJ, Alavanja MC. Pesticide use and colorectal cancer
593 risk in the Agricultural Health Study. *Int J Cancer* 2007;121(2):339-346. doi:10.1002/ijc.22635.
- 594 31. Reif J, Pearce N, Fraser J. Cancer risks in New Zealand farmers. *Int J Epidemiol* 1989;18(4):768-774.
595 doi:10.1093/ije/18.4.768.