ORIGINAL RESEARCH ARTICLE

Mapping research fronts of reproductive medicine using the NHANES database: A bibliometric and thematic analysis

DOI: 10.29063/ajrh2026/v30i6.7

*Jie Yang, Jun Liao, Yan Gao, Yue Gu,Yun-Xiang Chen, Ling-Yun Zhao, Qiu-Yi Yan and Hui-Min Li**

Guilin Medical University, Guilin, 541000, P. R. China

***For Correspondence:** Email: *huiminliwork@qq.com; lihuimin@glmu.edu.cn*: Phone*: +86 17377340573.*

Abstract

Reproductive health, infertility, and perinatal diseases significantly challenge human health development and continuity, requiring multidisciplinary approaches for enhanced academic understanding and practical solutions. This study used bibliometric analysis to examine related articles in the web of science (WoS) database, employing tools including bibliometrix and VOSviewer for systematic literature analysis. To capture the broader context of high-impact research, we first analyzed 1,026 highly-cited and hot papers in reproductive medicine, which revealed evolving thematic trends and global collaboration patterns. Then we examined 232 literatures using NHANES datasets for reproductive health research to identify research gaps and opportunities for further research in this field. The analysis revealed key trends, areas of concentration, and potential knowledge gaps in the literature, while the thematic analysis of reproductive health studies utilizing the NHANES database further underscored the wealth of data amenable to multidisciplinary investigation. Future research directions to explore new dimensions of reproductive health, improving health policies and clinical practices in this field using NHANES database are also suggested. The study concludes that multidisciplinary research using comprehensive datasets like NHANES is crucial for advancing understanding of reproductive health, infertility, and perinatal outcomes. *(Afr J Reprod Health 2026; 30 [7]: 66-79).*

Keywords: NHANES; Reproductive health; Infertility; Perinatal diseases; Bibliometrics

Résumé

La santé reproductive, l'infertilité et les maladies périnatales constituent un défi majeur pour le développement et la continuité de la santé humaine, nécessitant des approches multidisciplinaires afin d'améliorer la compréhension académique et de proposer des solutions pratiques. Cette étude a utilisé une analyse bibliométrique pour examiner les articles pertinents dans la base de données WoS, en employant des outils tels que Bibliometrix et VOSviewer pour une analyse systématique de la littérature. Afin de saisir le contexte plus large de la recherche à fort impact, nous avons d'abord analysé 1 026 articles hautement cités et à tendance chaude en médecine reproductive, ce qui a révélé les tendances thématiques évolutives et les modèles de collaboration internationale. Nous avons ensuite examiné 232 publications utilisant les jeux de données NHANES pour la recherche sur la santé reproductive, afin d'identifier les lacunes de recherche et les opportunités pour de futures études dans ce domaine. L'analyse a mis en évidence les tendances clés, les domaines de concentration et les lacunes de connaissances potentielles dans la littérature, tandis que l'analyse thématique des études sur la santé reproductive utilisant les jeux de données NHANES a en outre souligné la richesse des données susceptibles d'être étudiées de manière multidisciplinaire. Les orientations futures de la recherche visant à explorer de nouvelles dimensions de la santé reproductive, à améliorer les politiques sanitaires et les pratiques cliniques dans ce domaine à l'aide des jeux de données NHANES sont également proposées. L'étude conclut que la recherche multidisciplinaire utilisant des jeux de données complets tels que NHANES est cruciale pour faire progresser la compréhension de la santé reproductive, de l'infertilité et des issues périnatales. *(Afr J Reprod Health 2026; 30 [7]: 66-79).*

Mots-clés: NHANES; Santé reproductive; Infertilité; Maladies périnatales; Bibliométrie

Introduction

Reproductive medicine has been a rapidly growing and maturing field in medical science.[1] With economic development and increasing social pressures, reproductive health-related diseases have become a significant issue affecting human health and sustainable development.[2] The relationship between reproductive medicine and lifestyle,[3,4] environmental exposure,[5, 6] and systemic diseases,[7,

[8]has received widespread attention from researchers. Research on infertility can be mainly divided into basic experimental research, clinical research, and epidemiological research. Among these, epidemiological studies, with their large sample sizes and realistic research approach, are of significant value in uncovering the relationship between infertility and various lifestyles, environmental exposures, and systemic diseases. However, the relatively high requirements for sample size and sampling methods in epidemiological investigations have discouraged ordinary researchers due to insufficient human and material resources, thereby hindering the pace of uncovering the causes of infertility.

Therefore, comprehensive, standardized, and relatively open public databases with abundant data can provide a easy entry ticket for a large number of ordinary researchers.The National Health and Nutrition Examination Survey (NHANES) is a survey led by the Centers for Disease Control and Prevention (CDC) in the United States. NHANES employs a cross-sectional, population-based design aimed at collecting data on the diet, nutrition status, and health behaviors of the U.S. population. NHANES interviews cover demographic, socioeconomic, dietary, and health-related questions and provides researchers with a wealth of openly accessible research data. A comprehensive search of the Web of Science (WoS) database identified 232 publications to date that utilized NHANES data for reproductive health research. These studies investigated diverse determinants, including environmental exposures[5,6], dietary patterns[9,10], lifestyle factors[3,4], systemic diseases[11,12], and demographic and socioeconomic factors[2,13]. To contextualize our findings, we complemented this with an analysis of highly-cited and trending literature in the broader field of reproductive health. This approach allows for a critical examination of the capacity and limitations of the NHANES database in addressing contemporary research frontiers. Furthermore, this study conducts a bibliometric and thematic analysis of the NHANES-based literature and explores the database's relevant data modules, thereby proposing promising avenues for future research.

Methods

Data source and search strategy

A multi-level bibliometric information retrieval was conducted on the WoS platform by Clarivate Analytics. This study solely utilizes the WoS database for bibliometric analysis, capitalizing on its superior data quality, citation analysis capabilities, and extensive subject coverage to ensure the reliability and scientific validity of the research findings. The search was carried out in a single day on Nov. 12th, 2025, to prevent bias from daily updates of the open database.

The search query used for retrieving literatures in human reproductive health was "(TS=(infertility OR female genital disease OR female health conditions OR female infertility OR female reproductive disorder OR male genital disease OR male health condition OR male infertility OR newborn OR birth defect OR offspring health))". A total of 175355 records were retrieved. Only highly cited papers and hot papers in WoS were retrieved (1026 records) for further analysis.

Based on our previous preliminary literature pre-search results, the formal search query used for retrieving literatures concerning NHANES and reproductive health was "(TS=(infertility or female genital disease or female infertility or female reproductive disorder or male genital disease or male infertility or male genital disease or newborn or birth defect or offspring health)) AND TS=("NHANES") NOT TS=("KNHANES")". A total of 232 records were retrieved. Following manual screening, all included articles were confirmed to be studies on reproductive health based on the NHANES database, and therefore were fully retained.

Bibliometric analysis and visualization

In our study, the distribution of countries/regions, years of publication, and authors was analyzed by using the bibliometrix package in R (version 4.1.2). In brief, the plain text file containing the full record information of 1026/232 articles was imported into the bibliometrix work interface. Automatically, the

following analysis results were generated and then downloaded, which include the main information of the dataset, annual scientific production, graphical representation of most relevant sources, countries' scientific production.

The VOSviewer Version 1.6.20 (Centre for Science and Technology Studies, Leiden University, Leiden, Netherlands) was utilized to map distribution of keywords with more than 10 (for all reproductive health literatures)/2 (for literatures using NHANES database) co-occurrence into several clusters and color them by time course. Co-occurrence analysis identifies research hotspots and trends

Research key themes analysis

The exported Excel file, containing all the information of the 544 articles and the full texts of each article, was downloaded for thematic analysis of these published articles using the NHANES database. The primary factors under study and their correlations from each article were extracted, classified, and analyzed alongside the intervening factors.

All available data modules and research directions for study from NHANES (https://www.cdc.gov/nchs/nhanes/index.htm) official website across different cycles were downloaded and summarized. We compared the reproductive health-related themes already studied using NHANES data with the data modules available from NHANES, integrating our expertise in reproductive health to propose future directions for utilizing NHANES in related research. All the steps in the study were independently done by at least two researchers.

Ethics statement

This article is based on an analysis of publicly available literature databases and therefore does not require an ethics statement.

Results

Bibliometric analysis of highly cited papers in reproductive health

Based on a systematic bibliometric analysis of publications indexed in the WoS, a total of 1,026 documents published between 2015 and 2025 were retrieved and analyzed. The dataset comprises 628 articles, 388 reviews, 4 articles combined with proceedings papers, 3 articles combined with book chapters, 2 reviews in early access, and 1 review combined with a book chapter.

The annual scientific output exhibited fluctuations over the years, with publications ranging from a low of 78 in 2015 to a peak of 115 in 2022, and 52 documents already published in 2025 (data collection day). These publications were disseminated across 392 distinct sources. The most prolific journals included The Lancet (n = 54), New England Journal of Medicine (n = 27), Human Reproduction Update (n = 25), Nature (n = 25), and JAMA (n = 23). The journals spanned a wide range of high-impact domains, including general and internal medicine, public health, endocrinology, reproduction, and molecular biology.

Geographically, corresponding authors were predominantly affiliated with institutions in the United States (n = 4,036) and the United Kingdom (n = 1,448), followed by Iran (n = 1,032), Australia (n = 984), China (n = 886), and India (n = 723). Organizations such as Harvard University, the University of Oxford, and the University of Washington have made active contributions.

A detailed bibliometric analysis of keywords extracted from both titles and abstracts was performed using VOSviewer software. Keywords exceeding a minimum occurrence threshold were visualized to map the intellectual structure and evolution of the research field. Network visualization reveals a coherent and sophisticated clustering of terms, delineating several dominant research themes. These clusters can be categorized as follows: Cluster 1 encompasses themes of infertility mechanisms and reproductive biology, featuring terms like oocyte, follicle, pelvic pain, dysfunction, oxidative stress, cell, and vitro studies, with a focus on the pathological processes and cellular mechanisms underlying infertility, including explorations of obesity, placenta, and micropastic toxicity. Cluster 2 centers on infertility treatment and clinical reviews, highlighting concepts such as embryo, infertility treatment, surgery, reproduction, consensus, and miscarriage, emphasizing clinical practices, expert consensus, and therapeutic

interventions. Cluster 3 is dedicated to research methodologies and trial designs, including meta-analysis, controlled trial, confidence interval, bias, and neonatal mortality, focusing on the rigor of evidence generation through trials, data collection, and statistical analyses. Cluster 4 revolves around public health and pandemic-related reproductive research, with terms like pregnant woman, gestation, COVID-19, coronavirus disease, vaccine, and practice, investigating the impacts of pandemics and public health events on reproductive health. Cluster 5 addresses global disease burden and epidemiological indicators, encompassing country, global burden, disease study, disability, and life expectancy, focusing on population-level epidemiological analyses and health indicators across different regions.

Overlay visualization that maps keywords based on their average publication year clearly illustrates the temporal trajectory of research focus. Earlier research was predominantly anchored in foundational research methodologies and clinical trial frameworks, such as controlled trials, meta-analysis, and traditional clinical outcomes. In striking contrast, the more recent keywords, signify a pronounced shift towards cutting-edge reproductive biology mechanisms and global health perspectives. Contemporary research fronts are now dominated by in-depth investigations into cellular and molecular mechanisms of infertility (e.g., oxidative stress, cell dysfunction, micropastic toxicity) and a heightened focus on global health issues intersecting with reproduction, such as the impacts of pandemics and socioeconomic factors on reproductive outcomes. This evolution marks a transition from purely clinical and methodological explorations to a more integrated, mechanistic, and globally contextualized examination of reproductive health, reflecting the field's advancement toward precision medicine, interdisciplinary collaboration, and population health impact assessment, all hallmarks of high-impact SCI research.

Analysis of publications using NHANES data

Based on the systematic literature search conducted in the WoS, a total of 232 documents published between 2013 and 2025 were retrieved and curated for bibliometric analysis. The dataset is composed of 215 articles, 3 articles in early access, 5 meeting abstracts, 5 reviews, 2 editorial materials, 1 letter, and 1 correction. A comprehensive set of 507 author-defined keywords (DE) and 570 Keywords Plus (ID) terms was identified, reflecting the conceptual structure and thematic breadth of the research domain. Collaborative research was predominant, with a mean co-authorship rate of 5.44 per document and an international collaboration rate of 6.034%. In contrast to the collaborative norm, a small number of single-authored documents (n=1) were also published during this period.The annual scientific output demonstrated a marked upward trajectory, with the number of publications rising from 1 in 2013 to 87 in 2025, corresponding to an average annual growth rate of 45.09% (Figure 1A). These publications were disseminated across 94 distinct sources. The most prolific journals included Frontiers in Endocrinology (n=24), Frontiers in Nutrition (n=15), Scientific Reports (n=12), and BMC Public Health (n=11) (Figure 1B). The journals were categorized across multiple research domains, with significant representation in gynecology and reproductive health (e.g., BMC Women's Health, Reproductive Sciences), public health (Frontiers in Public Health, BMC Public Health), nutrition and metabolism (Frontiers in Nutrition, Nutrients), and endocrinology (Frontiers in Endocrinology, Reproductive Biology and Endocrinology).

Geographically, corresponding authors were primarily affiliated with institutions in China (n=446) and the United States (n=179), followed by South Korea (n=20), the United Kingdom (n=7), Iran (n=6), and Brazil (n=4) (Figure 1C). Organizations such as Jinan University, Guangzhou University of Chinese Medicine, and Harvard University have made active contributions.A detailed bibliometric analysis of keywords extracted from titles and abstracts reveals multiple dominant clusters with distinct color groupings (Figure 2A): One cluster revolves around the NHANES database and infertility research, featuring large nodes such as infertility, and pregnancy. It encompasses themes of epidemiological surveys (e.g., cross-sectional study, US population), reproductive health (e.g., female reproduction, fecundity), and associated outcomes like disease management and pregnancy-

related metrics, with a focus on leveraging NHANES data to explore reproductive health patterns and epidemiological trends. Another cluster centers on metabolic and obesity-related research, highlighted by prominent nodes like obesity, polycystic-ovary-syndrome, and insulin-resistance. It integrates concepts of metabolic disorders (e.g., cardiometabolic index, insulin resistance), adiposity (e.g., visceral obesity, body roundness index), and their intersections with reproductive conditions, investigating how metabolic imbalances and obesity influence fertility and reproductive health. A third cluster is dedicated to risk factors and psychological-epidemiological analysis, marked by nodes such as risk, depression, and anxiety. It involves psychological factors (e.g., anxiety, depression), demographic variables (e.g., United-States, adolescent, adults), and study design elements (e.g., epidemiology, validation, trends), focusing on the interplay between psychological states, demographic characteristics, and epidemiological methodologies in understanding health risks.

Additionally, there are clusters addressing lifestyle and environmental exposures (e.g., smoking, dietary, heavy metals, Bisphenol A) and biochemical and reproductive biomarkers (e.g., oxidative stress, folate, metabolites, oocyte). The former explores how lifestyle choices and environmental toxins impact health, while the latter delves into the biochemical markers and cellular mechanisms underlying reproductive processes, each with nodes of varying sizes reflecting their occurrence frequencies within the field.

Overlay visualization that maps keywords based on their average publication year clearly illustrates the temporal trajectory of research focus (Figure 2B). Earlier research (indicated by colors closer to blue/purple, spanning 2018–2020) was predominantly anchored in themes like methylmercury, birth-defects, and 25-hydroxyvitamin D, centering on environmental toxicants, birth defects, and vitamin D-related research. In striking contrast, more recent keywords (signified by colors leaning toward yellow, from 2022-2025) indicate a pronounced shift towards a sustained emphasis on infertility, and obesity, highlighting NHANES-based research, infertility, and the intersections of metabolic disorders and obesity. This evolution underscores a paradigm of shift from addressing acute environmental threats to managing complex, socio-metabolic determinants long-term well-being, with the rising research focus mirroring the growing public health challenge of infertility.

Scope and focus of reproductive health research using NHANES

Among the reproductive health literature published by NHANES, there is a lot related to "female infertility" (Table 1). Many researchers have studied various exposure factors associated with female infertility, such as heavy metals, organic compounds, pesticides (chemical exposure), as well as dietary fiber, nutritional intake, and polyunsaturated fatty acids (diet factors) impacting female infertility. Additionally, laboratory measurements, including BMI, waist circumference, body fat distribution, metabolic indicators, as well as inflammatory and insulin resistance markers, Chlamydia antibodies, and other blood measurements, are linked to female infertility. Factors such as lifestyle and socioeconomic status, including low income, disabilities, specific occupations, smoking, and sleep habits, also affect female infertility. Some studies incorporate mediating variables or interactive factors like vitamin D intake or depression scores to explore their combined effects with exposure factors on female infertility.

Beyond female infertility, NHANES has published a considerable amount of research on the relationship between various exposure factors and reproductive health-related conditions.

Figure 1: Output analysis of publications using NHANES data. (A) Publications production trend. (B) Most relevant journal categories. (C) Geographical distribution.

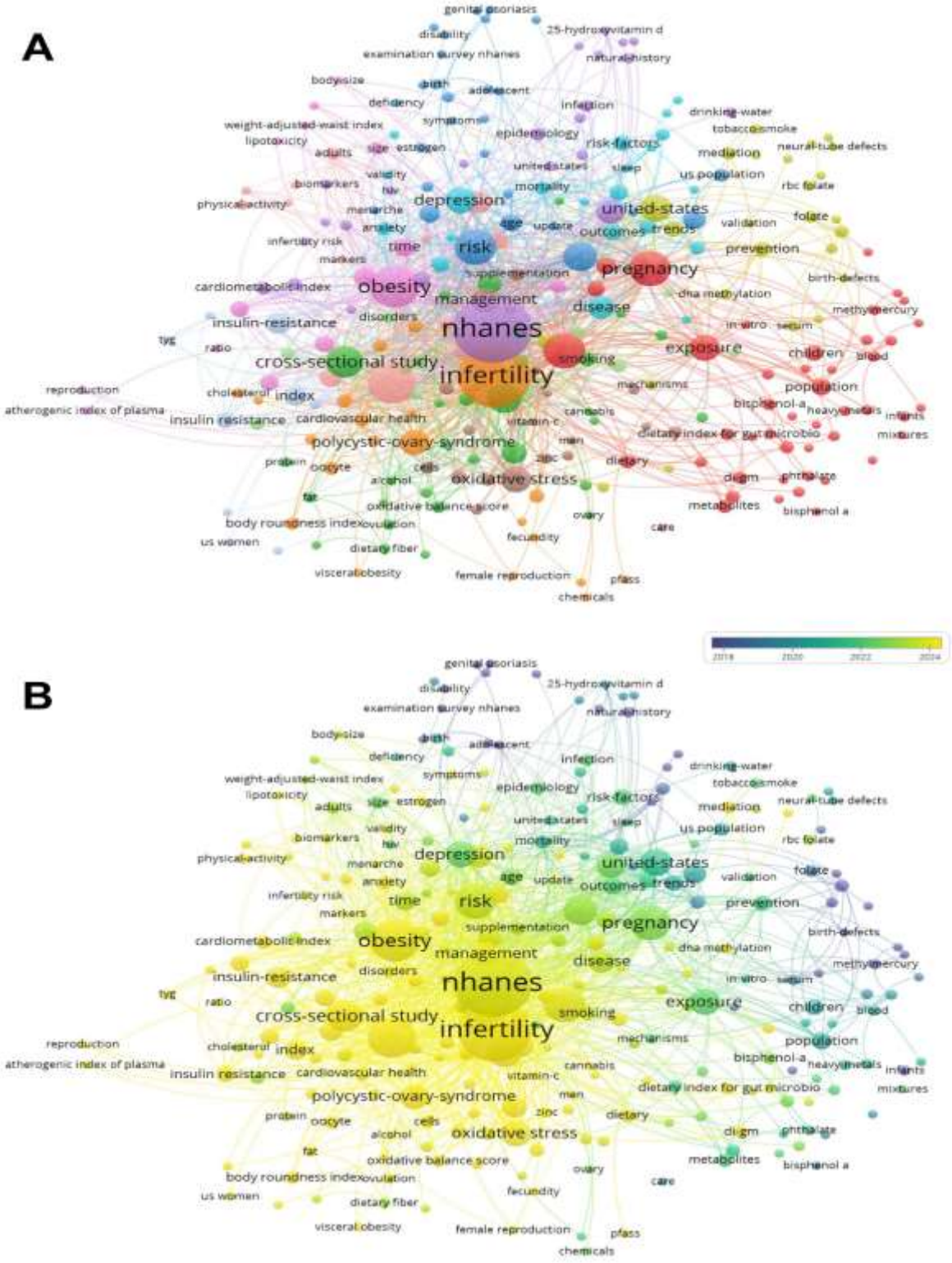

Figure 2: Co-occurrence analysis of publications using NHANES data. (A) Mapping of keywords of studies. (B) Overlay visualization of co-occurring author keywords by time (blue:earlier, yellow: later).

Table 1: Published articles concerning female infertility and varies exposure factors

Studied factor	Exposure factor	Type	Intervening factor
Female infertility	Pesticides exposure	Chemical exposure	Vitamin D Intake
Female infertility	Aldehydes exposure		/
Female infertility/Long-term amenorrhea	Cadmium, lead, mercury and mixtures		/
Female infertility	Co-exposure to phenols and phthalates		/
Female infertility	Di-2-ethylhexyl phthalate (DEHP)		/
Female infertility	Elevated blood mercury level		/
Female infertility	Endocrine disruptive chemicals (Parabens, Benzophenone-3, Bisphenol-A and Triclosan)		/
Female infertility	Heavy metals exposure		/
Female infertility	Perfluoroalkyl and Polyfluoroalkyl Substances		/
Female infertility	Single and multiple organophosphate pesticide exposure		/
Female infertility	Dietary fiber	Diet	/
Female infertility	Dietary inflammatory index		/
Female infertility	Nutrient intake		/
Female infertility	Polyunsaturated fatty acid		/
Female infertility	Body fat distribution	Metabolic index	/
Female infertility	Body mass index		/
Female infertility	Body shape index		/
Female infertility	Higher waist circumference		/
Female infertility	Metabolic healthy obesity		Depression score
Female infertility	Waist circumference		/
Female infertility	Weight-adjusted-waist index		/
Female infertility	Inflammatory markers	Blood test	/
Female infertility	Insulin resistance surrogates		/
Female infertility	Serum chlamydia trachomatis antibody levels		/
Female infertility	Serum uric acid		/
Female infertility	Disability	Lifestyle and socioeconomic situation	/
Female infertility	Lower socioeconomic status		/
Female infertility	Sleep behavior		/
Female infertility	Smoking status		/
Female infertility	Life's simple 7		/
Female infertility	HPV vaccination		/
Female infertility and other health problems	Flight attendant work		/

These studies explore reproductive health conditions such as reproductive system infections, reproductive endocrine disorders, pelvic floor dysfunction, reproductive system tumors, and sexual activity in relation to chemical exposure, infections, socioeconomic status, dietary patterns, and their associations with mental and various physiological diseases. Furthermore, research on offspring health has also been reported.

For instance, some scholars have focused on exploring the relationship of low birth weight and offspring overweight or diseases with maternal tobacco and environmental toxin exposure, as well as parental metabolic syndrome, providing epidemiological research foundations for the etiology of congenital or genetic diseases and several studies have utilized the NHANES database to investigate the exposure

Table 2: Variables related to reproductive health in NHANES data modules

First-level category	Second-level category	Third-level category
General data	Marital status	/
	Spouse's educational level	/
Examination	Reproductive system evaluation	/
	Sexually transmitted disease checking	/
Laboratory tests	STD tests	/
	TORCH panel	/
	Thyroid hormones	/
	Sexual hormons	/
	AMH	/
	Human epididymal secretory protein pM	/
	PSA	/
Life conditions and habits related	Nutrient intake	/
	Risk factor exposure	/
	Medication Usage	Female hormone pills
		Birth control pills
		Glucocorticoids
	Sexual activity	Number of sexual partners
		Modes of sexual activity
		Interest in sex
		Sexual frequency
		Ability to maintain an erection
Health condition and medical history	Gynecological issues	Uterine fibroid
		Endometriosis
		Infection
		Pelvic floor dysfunction
		History of gynecological operation
	Menstrual status and history	Age at the onset and cessation of menstruation
		Menstrual regularity
		Characteristics of menstrual flow
	Gestation and parturition status/history	Risk factors exposure (Nutrition, Tobacco)
		Pregnancy history (Current pregnancy)
		Live births
		Premature births
		Stillbirths
		Miscarriages
		Abortions
		Tubal pregnancies
		Infertility
		Maternal age at conception and delivery
		Times of vaginal and cesarean deliveries
		Birth weight of liveborn infants
		Breastfeeding history
		Postpartum menstrual resumption
		Postpartum depression
	Offspring health	Early childhood health status
		Adolescence health status
	Sexually transmitted diseases	/

rates or incidence rates of individual factors. For example, there has been research on the trends in urinary Bisphenol A concentrations among American adults, potentially impacting fertility. Other studied factors include infection rates of HPV, HSV, and Toxoplasma gondii among women of childbearing age, as well as indicators such as maternal nutrition adequacy during pregnancy and red blood cell folate concentrations among women of childbearing age, which can serve as references for foundational research on reproductive health and the formulation of related health policies.

Data modules in NHANES database

As a important program of the National Center for Health Statistics in the United States, NHANES encompasses over 400 unique topic modules and more than 14,000 variables. We categorizes the data modules and variables into “General Data”, “Health Conditions and Medical History”, “Physical Examination”, “Laboratory Tests”, and “Life Conditions and Related Habits” . “General Data” includes participant ID, gender, age, birthplace, nationality, ethnicity, language, education, family and marital status, and income. “Health Conditions and Medical History” cover areas like the circulatory, digestive, respiratory, urinary, endocrine, neurological, ophthalmological, cancer, infectious diseases, as well as male and female-specific conditions, and pregnancy-related complications. The database also includes physical examinations relevant to these systems in the “Physical Examination” section. In terms of “laboratory tests”, the database contains comprehensive data, including routine tests, biochemical tests, infection and immunity-related tests, toxins, and nutrient element analyses. Special tests such as bone density, lung function, and H. pylori tests are also included. The “Life Conditions and Related Habits” data incorporate dietary intake, consumption of beverages, alcohol and tobacco, medication use, screen time for TV and computers, duration of physical activity and outdoor time, frequency and duration of sexual activity, as well as experiences related to gun use and violence.

Variables related to reproductive health in NHANES data modules

As shown in Table 2, the "General Data" section includes variables closely associated with reproductive health such as marital status and the educational level of the spouse. In the "Health Condition and Medical History" category, relevant variables encompass gynecological issues, menstrual status and history, gestation and parturition status and history, offspring health, and sexually transmitted diseases. Gynecological problems include conditions such as uterine fibroids, endometriosis, a history of ovariectomy, pelvic infection, and pelvic floor dysfunction.Menstrual status and history cover factors like the age at menarche and menopause, menstrual regularity, and characteristics of menstrual flow. Gestation and parturition status/history and offspring health related variables include nutrition intake and risk factor exposure during pregnancy, pregnancy history (current pregnancy), live births, premature births, stillbirths, miscarriages, abortion, tubal pregnancies, infertility, maternal age at conception and delivery, the number of vaginal and cesarean deliveries, birth weight of live-born infants, breastfeeding history, postpartum menstrual resumption, postpartum depression, and the health status during the offspring's early childhood and adolescence.

“Examination” section comprises assessments of the reproductive system and screenings for sexually transmitted diseases. “Laboratory tests” category includes testing for STDs (sexual tramsmitted diseases), TORCH panel, thyroid hormones, sexual hormone levels, AMH, human epididymal secretory protein pM, and PSA. The "Life conditions and habits related" segment delves into nutrient intake like folic acid, tobacco use (particularly during pregnancy), medication use (such as female hormone pills, birth control pills, glucocorticoids), and sexual history. Sexual activity inquiry covers aspects like number of partners, sexual activities, sexual interest, frequency, and erectile function.

Reproductive and perinatal research prospects using NHANES

Given the vast and diverse content of the NHANES database with numerous variables, researchers can continue exploring uninvestigated variables related to their research areas for correlation or univariate analysis. For instance, future studies could delve into the relationships between female infertility, gynecological diseases, pregnancy outcomes and socioeconomic factors, living conditions, nutritional intake, toxic exposure, etc. As NHANES comprehensively records individual health data, it provides an opportunity to investigate the correlation between infertility and various internal/surgical/ENT(ear, nose, and throat) medical conditions, a current trend in dual-disease and multi-disease studies. Additionally, these data can be utilized to study the relationships among various gynecological disorders.

In the realm of male diseases, the database includes data on male hormone levels and erectile dysfunction, enabling research on male reproductive-related illnesses. Furthermore, NHANES records sexual activity data including sexual frequencies, modes of sexual intercourse, number of sexual partners, and sexually transmitted diseases, offering rich research materials for studying diseases in both genders. This can involve single-factor investigations or exploring relationships of reproductive diseases with other illnesses, social environments, lifestyles, and similar factors.In the realms of offspring health, the NHANES database also provides sufficient data variables, which include maternal folic acid intake, premature delivery, breastfeeding status, offspring birth weight, physiological indicators during childhood and adolescence. These variables can be utilized to explore the correlation between offspring health and parental exposure to factors such as environmental and nutritional conditions, gynecological and andrological disorders, various internal, surgical, and ENT conditions, infectious diseases, socioeconomic factors, and lifestyles.

Discussion

Mapping the frontier: hot topics in general reproductive health research

Our macroscopic analysis of the broader reproductive health literature reveals a field in the midst of a strategic evolution. The identified research fronts are increasingly polarized between two compelling directions: the deep dive into mechanistic biology and the expansive view of global health context. The intense focus on oxidative stress, cell dysfunction, and micropastic toxicity signifies a concerted effort to deconstruct infertility at the molecular and environmental levels, moving from symptomatic description to causal explanation. Concurrently, the strong emergence of clusters centered on public health, pandemic impacts (COVID-19), and global disease burden reflects a growing recognition that reproductive outcomes are inextricably linked to socioeconomic factors, healthcare systems, and global crises. High-impact journals consistently feature studies that either pioneer a novel mechanistic insight or offer a transformative, population-level perspective. This dual trajectory—towards both precision and context—sets the benchmark for what constitutes a cutting-edge contribution in the field, establishing a landscape where high-impact discoveries are driven by either biological novelty or global relevance.

The NHANES niche: strengths and limitations

The research paradigm utilizing the NHANES database has carved out a distinct and rapidly growing niche, defined by its unique capacity to bridge individual-level exposures and population-level health outcomes. Its primary strength in addressing contemporary reproductive health questions lies in its power to generate robust, generalizable hypotheses on modifiable risk factors. For instance, the database is exceptionally equipped to: Decipher Complex Comorbidities: Elucidate the intricate links between obesity, metabolic syndrome, and conditions like PCOS and female infertility on a large scale; Evaluate Lifelong and Transgenerational Impacts: Investigate how preconception nutritional status (dietary supplements, body mass index) or environmental exposures influence not only fertility but also pregnancy outcomes and long-term offspring health; Profile Real-World Patient Populations: Characterize the burden of depression, sleep disorders, and cardiovascular disease risk among individuals with reproductive conditions, providing

critical data for integrated care models; The observed surge in methodological sophistication—evidenced by the green cluster's cubic spline models, sensitivity analysis, and mediation analysis—represents the field's proactive response to its inherent limitations. Researchers are leveraging advanced statistics to approximate causal inference, control for a wide array of confounding factors, and model complex non-linear relationships, thereby extracting maximum insight from observational data.

Concerning female infertility research, researchers predominantly utilize the NHANES database to investigate the impacts of environmental exposure on female infertility,[14, 15] followed by considerations of nutrition[16, 17] and metabolic indicators.[18, 19] The NHANES database meticulously and comprehensively catalogs the participants' nutritional intake and toxicant exposure down to each micronutrient, toxin, various food types consumed, intake quantities, timings, frequencies, and even dining locations, providing researchers with a wealth of research materials. For instance, the modern trend of ordering takeout or dining out has been found to impact the onset of various diseases.[20] Could future researchers investigate the effects of cooking at home versus dining out on reproductive health? This could provide a reference for studying the impact of modern lifestyles on the current surge in infertility. Furthermore, researchers also explore the effects of sociodemographic factors such as occupation, income, and social status on infertility. Future researchers could refer to Table 2 and data modules within NHANES to select unexplored factors for further investigation.

Regarding male reproductive health, NHANES provides relatively fewer data modules and variables compared to females. Current research predominantly focuses on the relationship between sex hormones (primarily testosterone) and various toxins [21] and metabolic indicators.[11] In terms of functional indicators, researchers have investigated the association between erectile dysfunction and peripheral nerve dysfunction[7] or toxin exposure.[12] There is limited research on sexual orientation, sexual interest, frequency, and practices.

Potential research themes in reproductive and perinatal health utilizing NHANES

The NHANES database provides extensive modules and variables. By utilizing these data flexibly, a wide range of research topics can be explored. Currently, research on the impact of environmental exposures, nutritional intake, and lifestyles on reproductive health is relatively abundant, while studies on internal diseases, infectious diseases, and diseases from other specialties related to reproductive perinatology are limited. Also, NHANES is a continuously updated database. The newly released 2021-2023 NHANES cycle offers, for the first time, granular data on the COVID-19 pandemic—such as vaccination records, antibody serology, and prior infection—enabling a detailed examination of its broad implications for public health, with particular relevance to reproductive health outcomes.

In addition, dual-disease or multiple-disease studies are gradually becoming favored by researchers, focusing on exploring common pathogenic mechanisms or interactions between two or more diseases as an innovative approach to cross-sectional research. Future medical research is moving towards exploring big data, comprehensive health, and extensive interdisciplinary integration as its development direction. Furthermore, it is also a promising trend to study the relationship between various diseases in the human body. Just as reproductive health resembles a flower on a tree, it flourishes into its fullest potential when the human body, akin to a well-rooted, flourishing tree, is vigorous and healthy.

Limitations

There are some limitations to using the NHANES database for research. For instance, the NHANES database has relatively fewer molecular biology data, which is not conducive to the study of the mechanistic links between biological phenotypes and genetics. The data related to molecular biology research mainly include telomere length and the recently updated DNA methylation data. Compared with large databases with rich genetic data such as UKBiobank(UKB, https://www.ukbiobank.ac.uk/),

the NHANES database is difficult to conduct in-depth mechanism-related research. However, due to its completely free and easy-to-analyze without complex code, it is still a good choice for a large number of grassroots researchers. On the other hand, its US-centric nature limits its utility in addressing the global burden of reproductive disease in diverse genetic and cultural contexts, and its data structure is ill-suited for studying acute pandemic impacts in real-time.

Conclusion

Our study indicates that NHANES offers a wealth of accessible material for reproductive health-related research and provide future research directions for related research in this field. Recently, collaborative studies leveraging NHANES with other public databases and using multiple analysis methods such as Mendelian randomization, machine learning, as well as single-cell or multi-omics joint analysis have gain special interests of many researchers. Scientific researchers can utilized these large public databases to initially form epidemiological research basis. Then by incorporating their own clinical data with animal and cellular experimental results, researchers can establish a rigorous chain of evidence to substantiate their scientific hypotheses.

Acknowledgments

The authors acknowledge the financial support from the Guangxi Science and Technology Base and Talents Special Project (GuiKe AD22035022), the National Natural Fund Youth Project (82001621), Foundational Nursing Research Capacity Enhancement Project of Guilin Medical University (202504027), and the Graduate Education Innovation Program project 2024 (YCSW2024444).

Data availability statement

The data supporting this review's findings are from publicly accessible datasets and thus no additional source data need to be provided. The corresponding author is available to assist with data retrieval inquiries upon reasonable request.

Authors' contributions

Conceptualization: Hui-Min Li. Data curation: Jie Yang, Jun Liao. Formal analysis: Yue Gu, Qiu-Yi Yan, Ling-Yun Zhao. Funding acquisition: Hui-Min Li, Yue Gu. Project administration: Hui-Min Li. Writing-original draft: Jie Yang, Jun Liao, Hui-Min Li. Writing-review & editing: Hui-Min Li, Yan Gao, Yun-Xiang Chen, Qiu-Yi Yan, Ling-Yun Zhao. All authors mentioned in the article approved the manuscript.

Jie Yang and Jun Liao equally contributed to this study.

References

1. Nusinovich Y. Reproductive medicine. Science (New York, NY). 2023;380(6641):148-9.
2. Njagi P , Groot W. Financial costs of assisted reproductive technology for patients in low- and middle-income countries: a systematic review. Human reproduction open. 2023;2023(2):hoad007.
3. Wang L, Chang G, Cai S, Zou X, Qin M , Tan Y. The association of Life's Simple 7 and infertility among U.S. women. Front Endocrinol (Lausanne). 2024;15:1288289.
4. Smart SJ, Nikaj AN, Yu LL, Li HX, Yan F ,Zhang J. Association between maternal smoking during pregnancy and offspring overweight in US-born children. Pediatr Obes. 2021;16(3):9.
5. McClam M, Liu JH, Fan YH, Zhan TJ, Zhang Q, Porter DE, Scott GI ,Xiao S. Associations between exposure to cadmium, lead, mercury and mixtures and women's infertility and long-term amenorrhea. Arch PUblic Health. 2023;81(1):21.
6. Arya S, Dwivedi AK, Alvarado L, Kupesic-Plavsic S. Exposure of US population to endocrine disruptive chemicals (Parabens, Benzophenone-3, Bisphenol-A and Triclosan) and their associations with female infertility. Environ Pollut. 2020;265:8.
7. Hicks CW, Wang D, Windham BG , Selvin E. Association of Peripheral Neuropathy with Erectile Dysfunction in US Men. Am J Med. 2021;134(2):282-4.
8. Bowers K, Ehrlich S, Dolan LM, Gupta R, Altaye M, Ollberding NJ, Szczesniak R, Catalano P, Smith E, Khoury JC. Elevated Anthropometric and Metabolic Indicators among Young Adult Offspring of Mothers with Pregestational Diabetes: Early Results from the Transgenerational Effect on Adult Morbidity Study (the TEAM Study). J Diabetes Res. 2021;2021:10.
9. Bianchi CM, Mariotti F, Verger EO, Huneau JF. Pregnancy Requires Major Changes in the Quality of the Diet for Nutritional Adequacy: Simulations in the French and the United States Populations. PLoS One. 2016;11(3):17.
10. Orozco AM, Yeung LF, Guo J, Carriquiry A , Berry RJ. Characteristics of US Adults with Usual Daily Folic

Acid Intake above the Tolerable Upper Intake Level: National Health and Nutrition Examination Survey, 2003-2010. Nutrients. 2016;8(4):11.
11. Liu NZ, Luo XY, Li PY , Xiong W. The Triglycerides and Glucose Index is not superior to HOMA-IR in predicting testosterone deficiency among adult males. Andrology. 2023;11(2):215-24.
12. Glover F, Mehta A, Richardson M, Muncey W, Del Giudice F, Belladelli F, Seranio N ,Eisenberg ML. Investigating the prevalence of erectile dysfunction among men exposed to organophosphate insecticides. J Endocrinol Invest. 2023:11.
13. Chen XT, Liang JM, Yang Q, Huang JF, Li LX , Deng KX. Age affects the association between socioeconomic status and infertility: a cross-sectional study. BMC Womens Health. 2023;23(1):10.
14. Zhang WC, Cui YG , Liu JY. The association between aldehydes exposure and female infertility: A cross-sectional study from NHANES. Medicine. 2023;102(25):6.
15. Lin J, Lin XY, Qiu JH, You XM , Xu JB. Association between heavy metals exposure and infertility among American women aged 20-44 years: A cross-sectional analysis from 2013 to 2018 NHANES data. Front Public Health. 2023;11:9.
16. Chen MS, Wang DY, Gong HY, Zhang HM, Gao J and Luo SP. The association between dietary fiber and infertility among US women: The National Health and Nutrition Examination Survey, 2013-2018. Nutr Hosp. 2022;39(6):1333-40.
17. Ji XW, Ye Y, Wang L, Liu SY , Dong X. Association between nutrient intake and female infertility: a study based on NHANES database. J Obstet Gynaecol. 2023;43(2):10.
18. Zhong HX, Yu B, Zhao F, Cui HY, You LF, Feng D , Lu Y. Associations between weight-adjusted-waist index and infertility: Results from NHANES 2013 to 2020. Medicine. 2023;102(48):6.
19. Pan QW, Shen XL, Li HF, Zhu B, Chen DK , Pan JJ. Depression score mediate the association between a body shape index and infertility in overweight and obesity females, NHANES 2013-2018. BMC Womens Health. 2023;23(1):9.
20. Du Y, Rong S, Sun Y, Liu B, Wu Y, Snetselaar LG, Wallace RB , Bao W. Association Between Frequency of Eating Away-From-Home Meals and Risk of All-Cause and Cause-Specific Mortality. Journal of the Academy of Nutrition and Dietetics. 2021;121(9):1741-9.e1.
21. Meeker JD ,Ferguson KK. Urinary Phthalate Metabolites Are Associated With Decreased Serum Testosterone in Men, Women, and Children From NHANES 2011-2012. J Clin Endocrinol Metab. 2014;99(11):4346-52.