UC Davis
Dermatology Online Journal

Title

Free open access data sources for dermatology researchers

Permalink

Journal

Authors

Martin, Amylee
Liu, Jeffrey
Thatiparthi, Akshitha
et al.

Publication Date

2021

DOI

Copyright Information

Peer reviewed

Free open access data sources for dermatology researchers

Amylee Martin[1] BS, Jeffrey Liu[2] BS, Akshitha Thatiparthi[3] BS, Jashin J Wu[4] MD

Affiliations: [1]School of Medicine, University of California Riverside, California, USA, [2]Keck School of Medicine, University of Southern California Los Angeles, California, USA, [3]College of Osteopathic Medicine of the Pacific, Western University of Health Sciences, Pomona, California, USA, [4]Dermatology Research and Education Foundation, Irvine, California, USA

Corresponding Author: Jashin J Wu MD, Dermatology Research and Education Foundation, Irvine, CA, Email: jashinwu@hotmail.com

Abstract

Free open access data sources are a valuable tool for dermatology researchers and may uncover crucial information about dermatological diseases and delivery of dermatological care. This short review discusses six open access data sources including: National Health and Nutrition Examination Survey (NHANES), National Health Interview Survey (NHIS), National Survey of Children's Health (NSCH), National Ambulatory Medical Care Survey (NAMCS)/National Hospital Ambulatory Medical Care Survey (NHAMCS), Centers for Medicare and Medicaid Services (CMS), and Gene Expression Omnibus (GEO). We explain the role of each data source in dermatology and provide examples of past studies, which have used these data repositories.

Keywords: open access database, NHANES, NAMCS, GEO, NSCH, NHIS, CMS

Introduction

Free open access data sources provide insight into dermatological diseases by allowing analysis of large populations, without any monetary cost to the researcher and with minimal patient risk. Examination of these datasets by multiple investigators provides unique insights and generates more project ideas than could be produced by single researchers. This short review aims to inform dermatology researchers about six open access data sources to promote further use of these repositories to investigate skin disorders. These specific data sources are included, as all investigate one or more dermatological disorders, and each provides free, readily accessible datasets produced by a regulated organization. To provide context, we will discuss past studies that have utilized the data sources for dermatology research.

The following data sources will be discussed: National Health and Nutrition Examination Survey (NHANES), National Health Interview Survey (NHIS), National Survey of Children's Health (NSCH), National Ambulatory Medical Care Survey (NAMCS)/National Hospital Ambulatory Medical Care Survey (NHAMCS), Centers for Medicare and Medicaid Services (CMS), and Gene Expression Omnibus (GEO), (**Table 1**). The NHANES, NHIS, and NSCH are United States (U.S.) national surveys, which provide important epidemiologic information regarding chronic skin disease. The NAMCS extracts data from ambulatory care visits, whereas CMS extracts claims data from Medicare. Both provide insight into improving healthcare delivery. The last data source, GEO, contributes crucial genomic data and is useful for dermatology drug development.

Discussion

National Surveys Conducted by the National Center for Health Statistics (NCHS)

The National Center for Health Statistics (NCHS), operated by the Centers for Disease Control and Prevention, collects, analyzes, and disseminates data, which guides public policy and health promotion programs [1]. The NCHS conducts several national surveys [2]. We will discuss three, which are relevant to dermatology. These include NHANES, NHIS, and NAMCS/NHAMCS. The NSCH will be discussed in conjunction with the NHIS, since the results of these two surveys are often pooled together for analysis.

National Health and Nutrition Examination Survey (NHANES)
The NHANES is unique from the other NCHS surveys as it collects information from a physical examination, laboratory tests, and imaging in addition to questionnaire responses [3]. This survey addresses a wide array of diseases that may change each cycle. However, it emphasizes common chronic

Table 1. *Open data sources for dermatology researchers.*

Data Source	Description	Website	Role in Dermatology	Type of Data Collected
National Health and Nutrition Examination Survey (NHANES)	Annual U.S. based survey of non-institutionalized adults and children which assesses health and nutritional status	https://www.cdc.gov/nchs/nhanes/index.htm	Epidemiology and comorbidities of psoriasis, AD, hand dermatitis, skin cancer	-Interview which collects demographic, socioeconomic, dietary, and health-related information -Examination with medical, dental, and physiological measurements -Laboratory tests -Imaging studies
National Health Interview Survey (NHIS)	Annual U.S. based household interview survey of non-institutionalized adults and children which tracks health of the nation	https://www.cdc.gov/nchs/nhis/index.htm	Epidemiology and comorbidities of AD in children and adults	-Interview which collects information on household and each family unit residing in household -Interview of randomly chosen child and adult within each selected household which collects information on health status, healthcare services, and health-related behaviors
National Survey of Children's Health (NSCH)	U.S. based survey of non-institutionalized children ages 0-17 which provides insight into several areas of children's lives	https://www.childhealthdata.org/learn-about-the-nsch/NSCH	Epidemiology and comorbidities of AD in children	-Telephone interview or mail/web-based surveys with questions on physical/mental health, access to healthcare, family and social environment, neighborhood, and school
National Ambulatory Medical Care Survey/ National Hospital Ambulatory Medical Care Survey (NAMCS/NHAMCS)	Annual U.S. based survey of ambulatory medical care services	https://www.cdc.gov/nchs/ahcd/index.htm	Dermatology practice patterns	-Physician/staff member provided information on patient demographics, type of payer, reason for visit, visit diagnosis, medications, procedures, and other visit characteristics generated from office-based, outpatient department, and emergency department patient visits
Centers for Medicare and Medicaid Services (CMS)	Annual Medicare beneficiary claims data	https://data.cms.gov/	Dermatology disease prevalence and Medicare beneficiary utilization of dermatology services	-Medicare claims data for each calendar year including provider specialty, type of service, place of service, allowed services and charges, payment amounts by HCPC/CPT codes, utilization, and submitted charges
Gene Expression Omnibus (GEO)	Public genomics data repository of array- and sequenced-based data	https://www.ncbi.nlm.nih.gov/gds	Skin disorder gene expression profiles and disease pathogenesis	-Microarray -Next-generation sequencing -Other forms of high-throughput functional genomics data

disorders such as obesity, cardiovascular disease, and diabetes, as well as nutritional status and health behaviors [3]. Survey results, for each two-year cycle since 1999, are available for download on the website as SAS transport files [4,5]. Several software programs such as SAS, SUDAAN, SPSS, STATA, or R may be used to analyze the data files [6].

To accurately represent the U.S. population the NHANES uses a complex, multistage, probability sampling design that over-samples persons 60 and older, African Americans, and Hispanics [3,7]. In 2011, the NHANES introduced over-sampling of Asian Americans allowing for analysis of U.S. adults with origins from the Far East, South East, and the Indian subcontinent [8]. In addition, survey cycles with consistent sample designs may be appended to achieve larger sample sizes for subdomain analysis [4]. Lastly, the NHANES data must be analyzed with a "design-based" approach, which requires adjustments for differential clustering and stratification with sample weights [9].

The NHANES may be used to study prevalence and comorbidities of psoriasis, atopic dermatitis (AD), hand dermatitis, sun-exposure, and skin cancer. In most cycles, the NHANES includes a "dermatology questionnaire", which queries patients on sun-exposure [10]. Skin cancer is addressed in the "medical conditions questionnaire" from 1999-2016 and in the "dermatology questionnaire" from 1999-2004 [11]. Psoriasis can be found on the 2003-2006 "dermatology questionnaires." However, for the 2009-2014 cycles, psoriasis is located on the "medical conditions questionnaires" [11]. In addition to identifying patients with a self-reported psoriasis diagnosis, psoriasis severity was assessed in the 2003-2006 and 2011-2014 cycles [10]. The 2003-2004 cycle included psoriasis in the physical examination [12]. Atopic dermatitis was included on the 2005-2006 "allergy questionnaire" [13]. Hand dermatitis was part of the 2003-2004 dermatology physical examination, whereas "dermatitis, eczema, rash" items can be found on the 1999-2004 "dermatology questionnaires" [11,12].

Study Examples for NHANES
In 2009, Kurd et al. analyzed the 2003-2004 "dermatology questionnaire" and psoriasis physical

examination findings to determine the prevalence of psoriasis in the U.S. general population amongst diagnosed and undiagnosed individuals [14]. In 2020, Kao et al. estimated the prevalence of drug use in psoriasis patients from the 2009-2014 "drug use questionnaires" along with the "medical conditions questionnaires" to obtain psoriasis history [15]. Silverberg et al. utilized the 2005-2006 NHANES data from a uniaxial ActiGraph AM-7164 worn over a seven-day period to objectively assess physical activity in AD patients [16]. The 2003-2006 and 2010-2012 NHANES data was analyzed by Lai et al. to determine if sun-protective behaviors decline with time following a diagnosis of melanoma [17]. Using the data from the "dermatology questionnaire," Zamoiski et al. examined the correlation between self-reported sunscreen use and urinary concentration of 3-benzophenone (2003-2006 and 2009-2012), (**Table 2**), [18].

National Health Interview Survey (NHIS) and National Survey of Children's Health (NSCH)
The NHIS, a principal source for tracking information on the health of the US population, has been conducted yearly since 1957 [19]. This survey follows health trends, access to healthcare, and progress towards national health goals [19]. Sampled households are chosen using geographically clustered sampling techniques; however, the sample design is revised every decennial census [19]. The NHIS consists of a "core" questionnaire that is consistent with each cycle, as well as one or more questionnaires on "current" health topics [19]. The "core" questionnaire consists of four components: household, family, sample adult, and sample child [19]. The household and family questionnaires collect demographic information, whereas the sample adult and child questionnaires extract data on health status, health care services, and health behaviors [19]. The NHIS is administered in person by trained interviewers [20]. The data is available as ZIP files and may be analyzed with SAS, SPSS, and STATA software programs [21].

The NSCH is currently under the direction of Maternal and Child Health Bureau at the Health Resources and Services Administration [22]. The NCHS conducted the 2003, 2007, and 2011/2012

surveys via telephone [22]. The Census Bureau conducted the 2016-2018 surveys via the U.S. mail and internet [22]. Past surveys have addressed physical and mental health, access to healthcare, and environmental factors [22]. Cohorts were randomly selected, and the survey results were weighted to be representative of the U.S. population of non-institutionalized children ages 0-17 [22,23]. The 2016, 2017, and 2018 NSCH public-use files (PUF) are available for download on the Census Bureau's NSCH webpage as SAS and STATA data files. However, to access the 2003, 2007, and 2011/2012 data files, a "Request DRC Data Sets" form must be completed, which provides access to the datasets as SPSS, SAS, and STATA files [24,25].

The NHIS and NSCH may be used for AD research. Atopic dermatitis items are included on the NHIS 2010/2012 sample adult questionnaires [26]. The 2010 survey queries: "During the past 12 months, have you had dermatitis, eczema, or any other red, inflamed skin rash?" whereas the 2012 survey asks, "During the past 12 months, have you had eczema or any kind of skin allergy?" [27, 28]. Furthermore, the 2010 NHIS sample adult questionnaire includes 13 other AD questions, whereas the 2012 survey only includes one other dermatology question, which queries if the participant has had any other skin disorders besides AD in the last year [27, 28]. The 1997-2018 NHIS sample child questionnaires include the following AD question: "During the past 12 months, has 'sample child name' had eczema or any kind of skin allergy?" [29]. Lastly, AD item(s) are included on the 2003/2007 NSCH questionnaires [26]. The question which ascertains history of AD dermatitis is slightly different between the two survey years. Furthermore, the 2007 survey also queries about AD severity [30]. Since the NHIS and NSCH assess the prevalence of other diseases in addition to AD, these surveys are useful for studying comorbidities and risk factors for AD. Of note, the NHIS, NSCH, and NHANES data may be pooled together to allow for a larger sample size and more robust estimates.

Study Examples for NHIS and NSCH
In 2015, Silverberg et al. used the results from the 2008-2012 NHIS sample child questionnaires and the 2010/2012 NHIS sample adult questionnaires, combined with data from NCHS (2003/2007) and NHANES (2003-2006), to determine if AD is associated with short stature [26]. The height assessments were caregiver reported [26]. In 2016, Strom et al. investigated the link between AD and ADD/ADHD by combining data from the 1997-2013 NHIS sample child questionnaires and 2012 adult health questionnaire with NSCH findings (2003/2007), [31]. In 2010, Shaw et al. analyzed the results from the 2003 NSCH to determine the prevalence of AD in the U.S. [23]. They also examined the geographical distribution of the disease, comparing their results to epidemiologic studies conducted internationally [23]. Silverberg et al. pooled data from the 2005-2006 NHANES cycle with the 2010/2012 NHIS sample adult questionnaires to examine if individuals with adult AD are at increased risk for cardiovascular/cerebrovascular disease [32]. History of AD, coronary artery disease, angina, heart attack, stroke, and peripheral vascular disease were obtained from the questionnaire results [32]. In 2014, the odds ratio between AD and warts was calculated by Silverberg et al. using the 2007 NHIS sample child questionnaire (**Table 2**), [33].

National Ambulatory Medical Care Survey (NAMCS)/National Hospital Ambulatory Medical Care Survey (NHAMCS)
The NAMCS and NHAMCS, administered by the NCHS, provide practical insight into U.S. healthcare delivery [34]. The NAMCS has been conducted on a continuous annual basis since 1989 and extracts patient visit data from non-federally employed physicians working in ambulatory care outpatient settings [35,36]. Conducted annually since 1992, NHAMCS samples patient visits from non-federal emergency departments (ED) and hospital outpatient departments (OPD), [35,36].

Both surveys employ a complex, three-stage sampling method to capture 30 visits during the selected physician reporting week for NAMCS and 200 patient visits during the selected 4-week reporting period for NHAMCS [36]. The following information is collected from each sampled visit: patient demographics, type of payer, reason for visit, visit diagnosis (up to 3), medications, procedures,

other visit characteristics (i.e. counseling, screening, visit duration), and office characteristics (i.e. accepting new patients), [34]. Gilchrist et al. discovered that NAMCS under-reports behavioral counseling and over-reports visit duration [37]. After data collection, ICD-9 CM codes and "reason for visit" classifications are assigned [34]. The datasets are deidentified and made available, free of charge, on the website in SAS, SUDAN, SPSS, STATA formats [34].

Study Examples for NAMCS/NHAMCS
In 2014, Ahn et al. sought to better define the role of the NAMCS in dermatology. This study reported that NAMCS sampled a total of 29,554 dermatology visits from 1993-2010 [34]. McGregor et al. analyzed NAMCS/NHAMCS data (1993-2010) to characterize the demographic, geographic, and seasonal patterns associated with outpatient visits for AD [38]. This study also compared AD prescribing patterns between dermatologists and non-dermatologists [38]. In 2019, Vasicek et al. utilized the NAMCS (2003-2015), NHAMCS-OPD (2003-2011), and NHAMCS-ED (2003-2014) data to determine how frequently systematic corticosteroid and isotretinoin are prescribed together [39]. Al-Dabagh et al. used NAMCS (1989-2010) to determine the degree to which providers were following psoriasis treatment guidelines by examining how frequently psoriasis patients were prescribed systemic corticosteroids [40]. Reeder et al. utilized the NAMCS (1990-2009) data to examine wart treatment patterns and patient demographics (**Table 2**), [41].

Centers for Medicare and Medicaid Services (CMS)
The CMS oversees three major United States insurance programs including Medicare [42]. CMS has several publicly available datasets on the data.cms.gov website [43]. Limited Data Set files are not publicly available but may be requested [44].

Physician/Supplier Procedure Summary (PSPS) Master File
Dermatologists have utilized the PSPS Master File for research [45-48]. The PSPS data file summarizes the Medicare Part B carrier and durable medical equipment fee-for-service claims and is organized by carrier, pricing locality, Healthcare Common Procedure Coding System (HCPCS) code, HCPCS

modifier, provider specialty, type of service, and place of service [49]. The file is made available in July of each year and has been uploaded yearly since 2010 [49]. The datasets are available as a self-extracting ZIP CSV files and/or interactive web files and are compatible with software programs such as Microsoft Access, Microsoft SQL/DB2/Oracle, and SAS [49].

Part B National Summary Data File
The Part B National Summary Data File is another CMS data file that has been queried by dermatology researchers [47,50]. The files are summarized by Healthcare Common Procedure Coding/Current Procedural Terminology (HCPC/CPT) code ranges [51]. Each code range contains procedural, condition, and description subheadings, whereas each dataset displays allowed services, allowed charges, payment amounts by HCPC/CPT codes, and prominent modifiers [51]. The datasets are available in CSV format for the years 2000-2009 and in Excel format for 2010-2018 [51]. The data file can be analyzed with software programs such as Microsoft Excel, Microsoft Access, Microsoft SQL/DB2/Oracle, or SAS [51].

Physician and Other Supplier Public Use File
Dermatology researchers have also analyzed the Physician and Other Supplier Public Use File [50]. This data file provides valuable information regarding physician provided services and procedures received by Medicare beneficiaries [52]. The data file is organized by National Provider Identifier (NPI), Healthcare Common Procedure Coding System (HCPCS) code, and place of service. This data file provides information on utilization, payment (allowed amount and Medicare payment), and submitted charges [52]. Data for the years 2012 through 2017 may be downloaded from the website [52].

Study Examples for CMS Data Files
In 2013 Rogers et al. queried the PSPS Master File (1996-2008) to characterize skin cancer treatment trends [45]. Yeung et al. investigated Medicare patient utilization and associated costs for treatment of actinic keratosis by examining claims data from the PSPS Master File (2007-2015) with the Part B National Summary Data File (2007-2015), [47]. In

2015, Rogers et al. utilized the PSPS Master File to estimate the frequency of non-melanoma skin cancer in the Medicare population from 2006 to 2012 [46]. Using the PSPS Master File (2004-2009), Donaldson et al. determined the percentage of skin reconstruction (layered closures, grafts, and flaps) claims submitted to Medicare by dermatologists, compared to other specialties, using the 2004, 2007, and 2009 PSPS Master File [48]. To determine the frequency at which dermatologists bill for potassium hydroxide preps and ova and parasite examinations, Guzman et al. analyzed the Part B National Summary data file (2000-2016) and Physician and Other Supplier PUF (2012-2015), (**Table 2**), [50].

"Big Data"

Big Data is defined as having volume, velocity, and variety that exceeds traditional data processing capabilities [53]. These enormous datasets are hypothesis-generating and may allow for delivery of personalized medicine [53]. GEO, a large genomic dataset, is considered a "Big Data" source [53].

Gene Expression Omnibus repository (GEO)

In 2000, the National Center for Biotechnology Information (NCBI) created the Gene Expression Omnibus (GEO) to organize the large amount of data generated by microarray technologies [54]. The GEO is a free, publicly available, international genomic data repository that archives microarrays, next-generation sequencing, and other forms of high-throughput functional genomic data submitted by various researchers [55]. The website provides instructions to assist with accessing studies and gene expression patterns [56]. Platform records, sample records, and series records are submitted to GEO and compiled into "DataSets" and "gene Profile records." The GEO DataSets is a study-level database that is composed of GEO samples which are biologically comparable and statistically compatible. Gene Profiles is a gene-level database that links the expression of an individual gene across datasets [57].

The GEO has a powerful search engine, which allows users to identify data relevant to the topic of interest [58]. Original GEO records, datasets, and gene profiles may be downloaded. The website describes 6 different options for downloading the GEO records and 5 different options for downloading curated datasets and profiles [59]. Although intended for cystic fibrosis research, the comprehensive set of instructions provided by Guo et al. is applicable to other diseases [54].

Study Examples for GEO

To further elucidate the pathogenesis of psoriasis, Xie et al. examined gene expression pattern of psoriatic skin samples using the GSE14905 gene expression profile [60]. The GSE19743 gene expression profile was utilized by Xu et al. to determine key biomarkers associated with recovery of the skin following a severe burn [61]. Melanoma gene expression was examined by Zhao et al. using the GSE3189 gene expression profile [62]. Yuan et al. analyzed the GSE65127 gene expression profile to compare messenger RNAs and transcription factors in patients with vitiligo versus normal controls [63]. Using the GSE72702 gene expression profile, Teng et al. analyzed the gene expression pattern exhibited by hidradenitis suppurativa (**Table 2**), [64].

Conclusion

The six open access data repositories presented in this review may be utilized to investigate diverse dermatological topics and hypotheses. The NHANES is well suited to investigate chronic skin disorders as well as skin cancer. The NHIS and NSCH are useful for estimating AD disease prevalence and comorbidities, whereas the NAMCS/NHAMCS and CMS datasets provide insight into the delivery of dermatology care. The GEO provides knowledge of dermatological diseases at the molecular level.

Open access data sources may change the landscape of dermatology research. These data repositories allow researchers to perform quick analyses to explore promise of potential project ideas. Furthermore, eliminating the data collection step allows researchers to investigate study hypotheses more quickly and readily. Additionally, use of these existing data sources minimizes survey fatigue and reduces time burden for patients. Dermatology researchers would benefit from an increase in number of open access resources, as well as expansion of current open access data sources to include more dermatological diseases. In parting, the six open access data sources discussed in this review

have contributed greatly to the field of dermatology and their use should be encouraged and expanded.

Potential conflicts of interest

Dr. Wu is or has been an investigator, consultant, or speaker for AbbVie, Almirall, Amgen, Arcutis, Aristea Therapeutics, Boehringer Ingelheim, Bristol-Myers Squibb, Dermavant, Dr. Reddy's Laboratories, Eli Lilly, Galderma, Janssen, LEO Pharma, Mindera, Novartis, Regeneron, Sanofi Genzyme, Solius, Sun Pharmaceutical, UCB, Valeant Pharmaceuticals North America LLC, and Zerigo Health.

References

1. About NCHS - Celebrating 50 Years. https://www.cdc.gov/nchs/about/50th_anniversary.htm. Accessed on September 3, 2020.
2. CDC - NCHS - National Center for Health Statistics. https://www.cdc.gov/nchs/index.htm. Accessed on September 3, 2020.
3. NHANES - About the National Health and Nutrition Examination Survey. https://www.cdc.gov/nchs/nhanes/about_nhanes.htm. Accessed on September 3, 2020.
4. NHANES Frequently Asked Questions (FAQs). https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/faq.aspx. Accessed on September 19, 2020.
5. National Health and Nutrition Examination Survey (NHANES) | American Academy of Otolaryngology-Head and Neck Surgery. https://www.entnet.org/content/national-health-and-nutrition-examination-survey-nhanes. Accessed on September 19, 2020.
6. NHANES Tutorials - Module one - Datasets and Documentation. https://wwwn.cdc.gov/nchs/nhanes/tutorials/Module1.aspx. Accessed on September 3, 2020.
7. NHANES Tutorials - Module two - Sample Design. https://wwwn.cdc.gov/nchs/nhanes/tutorials/Module2.aspx. Accessed on September 3, 2020.
8. Paulose-Ram R, Burt V, Broitman L, Ahluwalia N. Overview of Asian American data collection, release, and analysis: National health and nutrition examination survey 2011-2018. *Am J Public Health.* 2017;107:916-921. [PMID: 28426300].
9. Center for Health Statistics N. *Vital and Health Statistics Series 2, Number 184 April 2020: Sample Design and Estimation Procedures Data Evaluation and Methods Research.*; 2015. https://www.cdc.gov/nchs/products/index.htm. Accessed on September 17, 2020.
10. NHANES Questionnaires, Datasets, and Related Documentation. https://wwwn.cdc.gov/nchs/nhanes/default.aspx. Accessed on September 3, 2020.
11. NHANES Search Variables. https://wwwn.cdc.gov/nchs/nhanes/search/default.aspx. Accessed on September 3, 2020.
12. NHANES 2003-2004: Dermatology Data Documentation, Codebook, and Frequencies. https://wwwn.cdc.gov/Nchs/Nhanes/2003-2004/DEX_C.htm. Accessed on September 3, 2020.
13. NHANES 2005-2006: Allergy Data Documentation, Codebook, and Frequencies. https://wwwn.cdc.gov/Nchs/Nhanes/2005-2006/AGQ_D.htm. Accessed on September 3, 2020.
14. Kurd SK, Gelfand JM. The prevalence of previously diagnosed and undiagnosed psoriasis in US adults: Results from NHANES 2003-2004. *J Am Acad Dermatol.* 2009;60:218-224. [PMID: 19022533].
15. Kao LT, Li IH, Wang WM, Lee HC, Kao HH, Pan KT. Illicit drugs, cannabis, and psoriasis in the United States: National Health and Nutrition Examination Survey. *J Am Acad Dermatol.* 2020;82:1514-1517. [PMID: 31923439].
16. Silverberg JI, Song J, Pinto D, et al. Atopic Dermatitis Is Associated with Less Physical Activity in US Adults. *J Invest Dermatol.* 2016;136:1714-1716. [PMID: 27189827].
17. Lai YC, Yew YW. Sunburns and Sun-Protective Behaviors after a Diagnosis of Melanoma. *Skinmed.* 2018;16:379-383. [PMID: 30575505].
18. Zamoiski RD, Cahoon EK, Michal Freedman D, Linet MS. Self-reported sunscreen use and urinary benzophenone-3 concentrations in the United States: NHANES 2003-2006 and 2009-2012. *Environ Res.* 2015;142:563-567. [PMID: 26298557].
19. NHIS - About the National Health Interview Survey. https://www.cdc.gov/nchs/nhis/about_nhis.htm. Accessed on September 3, 2020.
20. Strom MA, Silverberg JI. Utilization of Preventive Health Care in Adults and Children with Eczema. *Am J Prev Med.* 2016;50:e33-e44. [PMID: 26547540].
21. NHIS - 2018 Data Release. https://www.cdc.gov/nchs/nhis/nhis_2018_data_release.htm. Accessed on September 3, 2020.
22. National Survey of Children's Health - Data Resource Center for Child and Adolescent Health. https://www.childhealthdata.org/learn-about-the-nsch/NSCH. Accessed on September 4, 2020.
23. Shaw TE, Currie GP, Koudelka CW, Simpson EL. Eczema prevalence in the United States: Data from the 2003 national survey of children's health. *J Invest Dermatol.* 2011;131:67-73. [PMID: 20739951].
24. NSCH Data. https://www.census.gov/programs-surveys/nsch/data.html. Accessed on September 5, 2020.
25. Request an NSCH or NS-CSHCN Data Set--Data Resource Center. https://www.childhealthdata.org/help/dataset. Accessed on September 5, 2020.
26. Silverberg JI, Paller AS. Association between eczema and stature in 9 US population-based studies. *JAMA Dermatol.* 2015;151:401-409. [PMID: 25493447].
27. Index of /pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2010/English. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2010/English. Accessed on September 15, 2020.
28. Index of /pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2012/English. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2012/English. Accessed on September 15, 2020.
29. NHIS - 1997-2018. https://www.cdc.gov/nchs/nhis/1997-2018.htm. Accessed on September 15, 2020.

30. Archived Data Documents - NSCH and NS-CSHCN Prior to 2016. https://www.childhealthdata.org/learn-about-the-nsch/archive-prior-year-data-documents-and-resources. Accessed on September 5, 2020.

31. Strom MA, Fishbein AB, Paller AS, Silverberg JI. Association between atopic dermatitis and attention deficit hyperactivity disorder in U.S. children and adults. *Br J Dermatol.* 2016;175:920-929. [PMID: 27105659].

32. Silverberg JI. Association between adult atopic dermatitis, cardiovascular disease, and increased heart attacks in three population-based studies. *Allergy.* 2015;70:1300-1308. [PMID: 26148129].

33. Silverberg JI, Silverberg NB. Childhood atopic dermatitis and warts are associated with increased risk of infection: A US population-based study. *J Allergy Clin Immunol.* 2014;133:1041-1047. [PMID: 24094542].

34. Ahn CS, Allen M-M, Davis SA, Huang KE, Fleischer AB, Feldman SR. The National Ambulatory Medical Care Survey: A resource for understanding the outpatient dermatology treatment. *J Dermatolog Treat.* 2014;25:453-458. [PMID: 24256113].

35. NAMCS/NHAMCS - About the Ambulatory Health Care Surveys. https://www.cdc.gov/nchs/ahcd/about_ahcd.htm. Accessed on September 4, 2020.

36. Wu KK, Borba AJ, Deng PH, Armstrong AW. Association between atopic dermatitis and conjunctivitis in adults: a population-based study in the United States. *J Dermatolog Treat.* September 2019:1-5. [PMID: 31437060].

37. Gilchrist VJ, Stange KC, Flocke SA, McCord G, Bourguet C. A Comparison of the National Ambulatory Medical Care Survey (NAMCS) Measurement Approach With Direct Observation of Outpatient Visits. *Med Care.* 2004;42:276-280. [PMID: 15076827].

38. McGregor SP, Farhangian ME, Huang KE, Feldman SR. Treatment of atopic dermatitis in the United States: Analysis of data from the national ambulatory medical care survey. *J Drugs Dermatology.* 2017;16:250-255. [PMID: 28301621].

39. Vasicek B, Adams W, Steadman L, Reserva J, Swan J. Coprescription of isotretinoin and systemic corticosteroids for acne: An analysis of the national ambulatory medical care survey. *J Clin Aesthet Dermatol.* 2019;12:27-28. [PMID: 31360285].

40. Al-Dabagh A, Al-Dabagh R, Davis SA, et al. Systemic corticosteroids are frequently prescribed for psoriasis. *J Cutan Med Surg.* 2014;18:195-199. [PMID: 24800708].

41. Reeder VJ, Gustafson CJ, Davis SA, Fleischer AB, Huang WW. The treatment and demographics of warts: An analysis of national trends. *J Drugs Dermatology.* 2013;12:1411-1415. [PMID: 24301243].

42. CMS Homepage | CMS. https://www.cms.gov/. Accessed on September 5, 2020.

43. Data.CMS.gov | Data.CMS.gov. https://data.cms.gov/. Accessed on September 5, 2020.

44. Limited Data Set (LDS) Files | CMS. https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/DUA_-_NewLDS. Accessed on September 5, 2020.

45. Rogers HW, Coldiron BM. Analysis of skin cancer treatment and costs in the United States Medicare population, 1996-2008. *Dermatol Surg.* 2013;39:35-42. [PMID: 23199014].

46. Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the us population, 2012. *JAMA Dermatology.* 2015;151:1081-1086. [PMID: 25928283].

47. Yeung H, Baranowski ML, Swerlick RA, et al. Use and cost of actinic keratosis destruction in the Medicare part B fee-for-service population, 2007 to 2015. *JAMA Dermatology.* 2018;154:1281-1285. [PMID: 30326488].

48. Donaldson MR, Coldiron BM. Dermatologists perform the majority of cutaneous reconstructions in the Medicare population: Numbers and trends from 2004 to 2009. *J Am Acad Dermatol.* 2013;68:803-808. [PMID: 23453358].

49. Physician/Supplier Procedure Summary | CMS. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Physician-Supplier-Procedure-Summary. Accessed on September 5, 2020.

50. Guzman AK, Kaffenberger BH. Office-based dermatologic diagnostic procedure utilization in the United States Medicare population from 2000-2016. *Int J Dermatol.* 2019;58:1317-1322. [PMID: 31006848].

51. Part B National Summary Data File (Previously known as BESS) | CMS. https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Part-B-National-Summary-Data-File/Overview. Accessed on September 5, 2020.

52. Medicare Provider Utilization and Payment Data: Physician and Other Supplier | CMS. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/Physician-and-Other-Supplier. Accessed on September 5, 2020.

53. Wehner MR, Levandoski KA, Kulldorff M, Asgari MM. Research Techniques Made Simple: An Introduction to Use and Analysis of Big Data in Dermatology. *J Invest Dermatol.* 2017;137:e153-e158. [PMID: 28735616].

54. Guo Y, Townsend R, Tsoi LC. Simple Analysis of Deposited Gene Expression Datasets for the Non-Bioinformatician: How to Use GEO for Fibrosis Research. *Methods Mol Biol.* 2017;1627:511-525. [PMID: 28836221].

55. Frequently Asked Questions - GEO - NCBI. https://www.ncbi.nlm.nih.gov/geo/info/faq.html. Accessed on September 3, 2020.

56. About GEO DataSets - GEO - NCBI. https://www.ncbi.nlm.nih.gov/geo/info/datasets.html. Accessed on September 3, 2020.

57. GEO Overview - GEO - NCBI. https://www.ncbi.nlm.nih.gov/geo/info/overview.html. Accessed on September 3, 2020.

58. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: Archive for functional genomics data sets - Update. *Nucleic Acids Res.* 2013;41:D991-D995. [PMID: 23193258].

59. Download GEO data - GEO - NCBI. https://www.ncbi.nlm.nih.gov/geo/info/download.html. Accessed on September 3, 2020.

60. Xie S, Chen Z, Wang Q, Song X, Zhang L. Comparisons of gene expression in normal, lesional, and non-lesional psoriatic skin using DNA microarray techniques. *Int J Dermatol.* 2014;53:1213-1220. [PMID: 25041445].

61. Xu HT, Guo JC, Liu HZ, Jin WW. A Time-Series Analysis of Severe Burned Injury of Skin Gene Expression Profiles. *Cell Physiol Biochem.* 2018;49:1492-1498. [PMID: 30205399].

62. Zhao J, Zeng X, Song P, Wu X, Shi H. AKT1 as the PageRank hub gene is associated with melanoma and its functional annotation is highly related to the estrogen signaling pathway that may regulate the growth of melanoma. *Oncol Rep.* 2016;36:2087-2093. [PMID: 27573172].

63. Yuan X, Meng D, Cao P, et al. Identification of pathogenic genes and transcription factors in vitiligo. *Dermatol Ther.* 2019;32. [PMID: 31306558].

64. Teng Y, Tao X, Lu W, et al. Identification of Hidradenitis Suppurativa-Related mRNA Expression Patterns Through Analysis

of Gene Expression Omnibus. *Dose-Response*. 2020;18. [PMID: 32782449].

Table 2. *Example studies using National Health and Nutrition Examination Survey (NHANES), National Health Interview Survey (NHIS), National Survey of Children's Health (NSCH), National Ambulatory Medical Care Survey (NAMCS)/National Hospital Ambulatory Medical Care Survey (NHAMCS), Centers for Medicare and Medicaid Services (CMS), and Gene Expression Omnibus (GEO).*

Title	Author (year)	Data Source	Cycles/Gene Expression Profile Used	Topic Investigated
The prevalence of previously diagnosed and undiagnosed psoriasis in US adults: Results from NHANES 2003-2004	Kurd (2009), [14]	NHANES	2003-2004	Psoriasis prevalence
Illicit drugs, cannabis, and psoriasis in the United States: National Health and Nutrition Examination Survey	Kao (2020), [15]	NHANES	2009-2014	Illicit drug use in psoriasis patients
Atopic Dermatitis Is Associated with Less Physical Activity in US Adults	Silverberg (2016), [16]	NHANES	2005-2006	Physical activity in AD patients
Sunburns and Sun-Protective Behaviors after a Diagnosis of Melanoma	Lai (2018), [17]	NHANES	2003-2006; 2010-2012	Sun-protective behaviors in patients with a history of melanoma
Self-reported sunscreen use and urinary benzophenone-3 concentrations in the United States: NHANES 2003-2006 and 2009-2012	Zamoiski (2015), [18]	NHANES	2003-2006; 2009-2012	Sunscreen use and urinary benzophenone-3 concentration
Association between eczema and stature in 9 US population-based studies	Silverberg (2015), [26]	NHIS NHANES NSCH	2008-2012 sample child questionnaire; 2010/2012 sample adult questionnaire 2003-2006 2003/2007	Short stature and AD in adults and children
Association between atopic dermatitis and attention deficit hyperactivity disorder in U.S. children and adults	Strom (2016), [31]	NHIS NSCH	1997-2013 sample child questionnaire; 2012 sample adult questionnaire 2003/2007	AD and ADD/ADHD in adults and children
Eczema prevalence in the United States: Data from the 2003 national survey of children's health	Shaw (2010), [23]	NSCH	2003	AD prevalence in children and geographic distribution
Association between adult atopic dermatitis, cardiovascular disease, and increased heart attacks in three population-based studies	Silverberg (2015), [32]	NHANES NHIS	2005-2006 2010/2012 sample adult questionnaire	Risk of cardiovascular and cerebrovascular disease in adults with AD
Childhood atopic dermatitis and warts are associated with increased risk of infection: A US population-based study	Silverberg (2014), [33]	NHIS	2007 sample child questionnaire	Association between AD and warts
The National Ambulatory Medical Care Survey: A resource for understanding the outpatient dermatology treatment	Ahn (2014), [34]	NAMCS	N/A	Role of NAMCS in dermatology

Treatment of atopic dermatitis in the United States: Analysis of data from the national ambulatory medical care survey	McGregor (2017), [38]	NAMCS/ NHAMCS	1993-2010	Demographic, geographic, and seasonal patterns associated with outpatient visits for AD; prescribing patterns for AD dermatologists versus non-dermatologists
Coprescription of isotretinoin and systemic corticosteroids for acne: An analysis of the national ambulatory medical care survey	Vasicek (2019), [39]	NAMCS	2003-2015	Frequency of prescribing systematic corticosteroid with isotretinoin
		NHAMCS-OPD	2003-2011	
		NHAMCS-ED	2003-2014	
Systemic corticosteroids are frequently prescribed for psoriasis	Al-Dabagh (2014), [40]	NAMCS	1989-2010	Frequency of which psoriasis patients were prescribed systemic steroids
The treatment and demographics of warts: An analysis of national trends	Reeder (2013), [41]	NAMCS	1990-2009	Wart treatment patterns and patient demographics
Analysis of skin cancer treatment and costs in the United States Medicare population, 1996-2008	Rogers (2013), [45]	PSPS Master File	1996-2008	Skin cancer treatment trends in the Medicare population
Use and cost of actinic keratosis destruction in the Medicare part B fee-for-service population, 2007 to 2015	Yeung (2018), [47]	PSPS Master File Part B National Summary Data file	2007-2015 2007-2015	Medicare patient utilization and associated costs for treatment of actinic keratosis
Incidence Estimate of Nonmelanoma Skin Cancer (keratinocyte carcinomas) in the US Population, 2012	Rogers (2015), [46]	PSPS Master File	2006-2012	Frequency of non-melanoma skin cancer in the Medicare population
Dermatologists perform the majority of cutaneous reconstructions in the Medicare population: Numbers and trends from 2004 to 2009	Donaldson (2013), [48]	PSPS Master File	2004, 2007, 2009	Percentage of skin reconstruction (layered closures, grafts, and flaps) claims submitted to Medicare by dermatologists relative to other specialties
Office-based dermatologic diagnostic procedure utilization in the United States Medicare population from 2000-2016	Guzman (2019), [50]	Part B National Summary data file Physician and Other Supplier PUF	2000-2016 2012-2015	Frequency at which dermatologists bill Medicare for potassium hydroxide preps, and ova and parasite examinations
Comparisons of gene expression in normal, lesional, and non-lesional psoriatic skin using DNA microarray techniques	Xie (2014), [60]	GEO	GSE14905 gene expression profile	Gene expression profile of psoriatic skin samples
A Time-Series Analysis of Severe Burned Injury of Skin Gene Expression Profiles	Xu (2018), [61]	GEO	GSE19743 gene expression profile	Key biomarkers associated with recovery of the skin following a severe burn
AKT1 as the PageRank hub gene is associated with melanoma and its functional annotation is highly related	Zhao (2016), [62]	GEO	GSE3189 gene expression profile	Melanoma gene expression

to the estrogen signaling pathway that may regulate the growth of melanoma		- 11 -		
Identification of pathogenic genes and transcription factors in vitiligo	Yuan (2019), [63]	GEO	GSE65127 gene expression profile	Messenger RNAs and transcription factors in patients with vitiligo versus normal controls
Identification of Hidradenitis Suppurativa-Related mRNA Expression Patterns Through Analysis of Gene Expression Omnibus	Teng (2020), [64]	GEO	GSE72702 gene expression profile	Gene expression pattern exhibited by hidradenitis suppurativa