UCONN
UNIVERSITY OF CONNECTICUT

UCONN
CENTER FOR MHEALTH AND SOCIAL MEDIA

Large national datasets assessing health-related use of the internet, technology, and social media

Molly E. Waring, PhD
Assistant Professor, Department of Allied Health Sciences
Director of Methodology Core, UConn Center for mHealth and Social Media
Email molly.waring@uconn.edu
Twitter @DrMollyWaring
6 April 2020

Sampling from a population

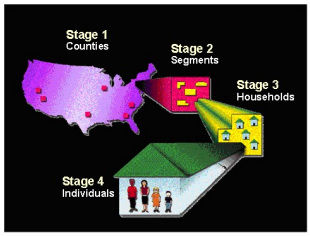
- To answer research questions, ideally we would observe all members of the population of interest
- Since this isn't feasible (practically, ethically, financially), we study a representative sample of the population
  - Sampling unit = entity selected for study
  - e.g., individual, dyad, household, medical visit
- With statistical inference, conclusions based on the sample can be attributed back to the population

Sample weights: the basic idea

- The statistical weight of a sampled person is the number of people in the population that this person represents
- If we surveyed all members of the population
  - Each person represents 1 person
  - Each person has a weight of 1/1 = 1
- Simple random sample with sampling rate = 1/200
  - Each sampled person represents 200 people
  - Each sampled person has a weight of 1/200

More complex sampling

- In more complex sample
  - Sampling weight = probability of selection
    - E.g., P(select county)*P(select segment)*P(select household)*P(select individual)
  - Each sampled person represents 1/weight people
- Sampling weights may also be adjusted for non-response or adjusted to represent population characteristics



https://www.cdc.gov/nchs/tutorials/NHANES/SurveyDesign/SampleDesign/Info1.htm

Weighting analyses to generate estimates representative of the population

- Using analytic techniques that incorporate sample weights allow us to generate results that are representative of the underlying population
- Results from appropriately weighted analyses can be nationally-representative
  - E.g., "x% of US adults"
- Results are estimates, report SE or 95% CI
  - E.g., "50.7% (SE: 2.5%) of US adults…"
  - Different from 50.7% of a sample (exact percent)

Ethical considerations

- IRB review required?
  - For original data collection, absolutely.
  - For research using the de-identified, public-use datasets, no.
  - For research using limited-access datasets that include identifiers or sensitive data, typically yes.

Overview of existing datasets

- Adults: HINTS, Pew, NHIS, CPS
- Children/adolescents: FLASHE, YRBS, NSCH
- Posts/searches: Twitter, Google trends

- https://hints.cancer.gov
- HINTS was designed to assess the impact of the health information environment
  - How people access and use health information
  - The degree to which people are engaged in healthy behaviors
- American public's access to and use of info about cancer from prevention, early detection, diagnosis, treatment, and survivorship
  - Including cancer risk perceptions, communication, risk behaviors, understanding of cancer prevention messages

Timeframe

- HINTS 1 = 2003 (n=6369)
- HINTS 2 = 2005 (n=5586)
- HINTS 3 = 2008 (n=7674)
- HINTS Puerto Rico = 2009 (n=639)
- HINTS 4
  - Cycle 1 = 2011 (n=3959)
  - Cycle 2 = 2012 (n=3630)
  - Cycle 3 = 2013 (n=3185)
  - Cycle 4 = 2014 (n=3677)
- HINTS FDA
  - Cycle 1 = 2015 (n=3787)
  - Cycle 2 = 2017 (n=1736)
- HINTS 5
  - Cycle 1 = 2017 (n=3285)
  - Cycle 2 = 2018 (n=3504)
  - Cycle 3 = 2019 (n=5438)
  - Cycle 4 = 2020 (planned)

https://hints.cancer.gov/about-hints/data-collection-schedule.aspx

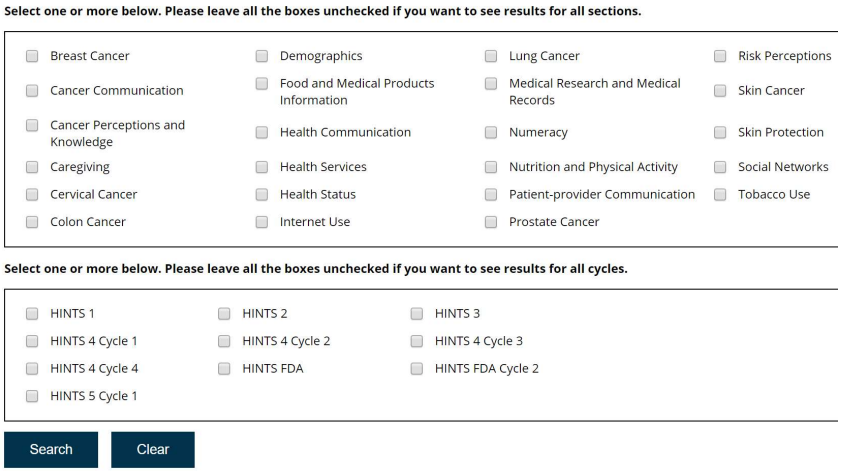
Design, Sampling, & Generalizability

- HINTS is a mail survey based on a stratified probability sample of the US adult, civilian, non-institutionalized population
  - Stratification: addresses in areas with high vs low concentrations of minority populations; oversample high-minority areas
  - Randomly selected addresses, adult within households
  - A Spanish version of the questionnaire was distributed to households that had a Hispanic surname
- Weighted analyses are generalizable to Civilian, non-institutionalized US adults
  - Can compare rural vs. urban metropolitan statistical areas (MSAs), census regions (4), and census divisions (9)

https://hints.cancer.gov/docs/methodologyreports/HINTS5_Cycle_2_Methodology_Report.pdf

Data collection

- Content varies year to year based on trending areas/recent developments in cancer communication
- Ex: HINTS 5 Cycle 2 includes questions on caregiving, palliative care, and family cancer history



Search questions/topics: https://hints.cancer.gov/view-questions-topics/all-hints-questions.aspx
Survey instruments: https://hints.cancer.gov/data/survey-instruments.aspx

Questions of interest available in HINTS 5 Cycle 2 (2018) and Cycle 3 (2019)

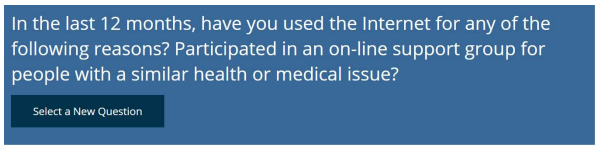
- Internet use; how accessed (e.g., wifi, broadband; home, work)
- Online activities: online support group, watch health-related video on YouTube, social networking site like Facebook or LinkedIn, share health-related info on social media sites, communicate with doctor, seek health information
- Info-seeking: sources, trust in sources
- Have health-related apps
- Use of electronic monitor/device to track health or activity such as Fitbit, BP monitor, blood glucose monitor; willingness to share info with doctor; willingness to share info with family or friends

Example studies

- Swoboda2019: online health information-seeking and meeting cancer prevention behavior recommendations (F&V intake, exercise, smoking, mammogram, Pap test, colon cancer screening) (HINTS 2007, 2011, 2012, 2014, 2017) https://link.springer.com/article/10.1007%2Fs13187-019-01597-0
- Alcala2019: Trust in Sources of Tobacco Health Information, Perceptions of Harm, and Use of E-Cigarettes https://www.ncbi.nlm.nih.gov/pubmed/30715455
- Bangerter2019: Health Information–Seeking Behaviors of Family Caregivers https://aging.jmir.org/2019/1/e11237/
- Greenberg-Worisek2019: Tracking Healthy People 2020 Internet, Broadband, and Mobile Device Access Goals: An Update Using Data from HINTS https://www.jmir.org/2019/6/e13300/
- Mahmood2019: Use of mobile health applications for health-promoting behavior among individuals with chronic medical conditions https://www.ncbi.nlm.nih.gov/pubmed/31656632

Accessing HINTS data – quick stats

Green = cycles that included the question – click on cycle to bring up responses



		ESTIMATED US ADULT POPULATION		SURVEY RESPONDENTS	
	Response	Number	Percentage	Responses	Percentage
1	Yes	13,863,167	5.6	203	5.8
2	No	229,458,262	92	3221	91.9
-9	Missing data (Not Ascertained)	6,168,343	2.5	80	2.3
	Total	-	**100%**	**3,504**	**100%**

Sample n (%) and weighted N (%) – useful for assessing feasibility of study ideas and quick statistics

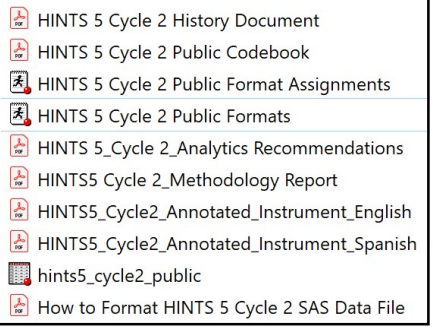
https://hints.cancer.gov/view-questions-topics/all-hints-questions.aspx

Downloading HINTS datasets

- https://hints.cancer.gov/data/download-data.aspx
- Must agree to terms of use
- Zip file includes data, formats, and documentation including survey annotated with skip patterns
- SAS, Stata, and SPSS versions available

B3. In the past 12 months, have you used the Internet to look for information about cancer for yourself? InternetCancerInfoSelf

[1] Yes
[2] No



Strengths & Limitations

- Data on cancer-related risk perceptions, communication, information-seeking, knowledge, and screening/prevention
- No/limited data on other health topics
- Can examine trends across time
- Questions asked vary across cycles
- Cross-sectional
- Data self-reported
- Low response rate (32% in 2017, 33% in 2018)
  - Sample weights do adjust for non-response

Pew Research Center

- https://www.pewresearch.org/internet/
- Internet, science, & tech research – focus on how science and technology changes affect families, communities, education, health care and medicine, civic and political life, and workers' activities
  - Plus lots of other topics! https://www.pewresearch.org/topics/
- RDD sample of landlines and cell phones
- Results representative of US adult population

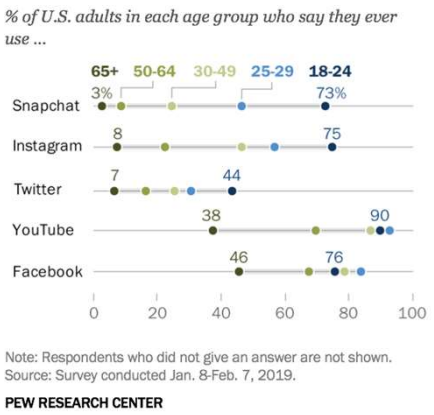
https://www.pewresearch.org/methods/u-s-survey-research/our-survey-methodology-in-detail/
https://www.pewresearch.org/methods/u-s-survey-research/american-trends-panel/

Questions of Interest

- Use of social media platforms, frequency of use
- Broadband at home vs smartphone access
- Ownership of desktop computers, laptops, tablets, e-readers, smartphones, cell phones
- Digital knowledge
- Perceptions of Facebook policies, cybersecurity
- Thoughts about future of technology
- … and many other questions related to use of the Internet, technology, and social media

Accessing data

- https://www.pewresearch.org/internet/
- Fact Sheets
  - E.g., Mobile, Social Media, Internet/Broadband
- Reports
  - E.g., Americans and Digital Knowledge, Virtues and Downsides of Online Dating
- Most survey datasets available to download
  - Need to create account



https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/

Considerations

- Focus on prevalence of activities and differences by demographics (age, gender, race/ethnicity, education, income, rural/suburban/urban)
- Reports examine unadjusted differences between groups
- Typically little other health information collected

National Health Interview Survey (NHIS)

- https://www.cdc.gov/nchs/nhis/index.htm
- Nation's largest in-person household health survey (2018: 25,417 adults and 8,269 children)
- Conducted continuously since 1957
- Collects information on health status and conditions, disability, access to and use of health services, health insurance coverage, immunizations, risk factors, and health-related behaviors
- Generalizable to civilian noninstitutionalized US population

Core Questionnaires

Family File
- I. Coverage Section: Telephone (COV)
- II. Family Structure Variables (MFM)
- III. Family Food Security Section (FFS)

Person File
- I. Household Composition Section (HHC)
- II. Family Identification Section (FID)
- III. Health Status and Limitation of Activity Section (FHS)
- IV. Health Care Access and Utilization Section (FAU)
- V. Health Insurance Section (FHI)
- VI. Socio-demographic Section (FSD)
- VII. Income and Assets Section (FIN)
- VIII. English Language Proficiency Section (FLG)

Sample Child File
- I. Child Identification Section (CID)
- II. Child Conditions, Limitation of Activity and Health Status Section (CHS)
- III. Child Health Care Access and Utilization Section (CAU)
- IV. Child Mental Health Brief Supplement (CMB)
- V. Child Influenza Immunization Supplement (CFI)

Sample Adult File
- I. Adult Identification Section (AID)
- II. Adult Socio-Demographics Section (ASD)
- III. Adult Conditions Section (ACN)
- IV. Adult Health Status and Limitation of Activity Section (AHS)
- V. Adult Health Behaviors Section (AHB)
- VI. Adult Health Care Access and Utilization Section (AAU)
- VII. Adult Selected Items Section (ASI)
- VIII. Adult Cancer Screening (NAF)
- IX. Adult Functioning and Disability Supplement File (AFD)
- X. Adult Internet and Email Usage Section (AWB)

*** partial table

Table 10. Conditions and reference periods in the core 2018 National Health Interview S

Question number	Condition	Ever	12 months
ACN.010	Hypertension	X	
ACN.020	Hypertension 2+ visits	X	
ACN.031	Coronary heart disease	X	
ACN.031	Angina pectoris	X	
ACN.031	Heart attack (MI)	X	
ACN.031	Other heart condition or heart disease	X	
ACN.031	Stroke	X	
ACN.031	Emphysema	X	
ACN.035	Chronic obstructive pulmonary disease (COPD)	X	
ACN.080	Asthma	X	
ACN.085	Asthma still have		

- Cigarette smoking
- Leisure-time PA
- Alcohol use
- Height & weight

ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2018/srvydesc.pdf

For survey instruments: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2018/english/

Supplemental Surveys Questions of Interest

2017
- Complementary Health
- Cultural Competence (CLAS)
- Heart Disease and Stroke
- Cognitive Disability
- Epilepsy
- Diabetes
- Vision
- Chronic Pain
- Hepatitis
- Expanded Content on Health Care Access and Utilization
- Food Security
- Child Mental Health (Brief Strengths and Difficulties Questionnaire)
- Immunization
- Heart Disease and Stroke Prevention (Million Hearts)
- Tobacco and E-cigarette Use
- Internet and Email Usage
- Disability

2018
- Asthma
- Cancer Screening
- Cognitive Disability
- Food Security
- Immunization
- Heart Disease and Stroke Prevention (Million Hearts)
- Tobacco and E-cigarette Use
- Internet and Email Usage
- Disability

2018 Adult Questionnaire

- Use Internet, email
- Used computer to look online for health information, fill prescription, schedule medical appointment, communicate with health care provider

*** different questions each year

https://www.cdc.gov/nchs/data/nhis/NHIS_Supplements_and_Sponsors.pdf
ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Survey_Questionnaires/NHIS/2018/english/

Example studies

- Hung2020: Health information technology use among older adults in the US (NHIS 2009-2018) https://www.ncbi.nlm.nih.gov/pubmed/32096650
- Kindratt2019: Health information technology use and influenza vaccine uptake among US adults (NHIS 2011-2015) https://www.ncbi.nlm.nih.gov/pubmed/31445279
- Gonzalez2019: Web-Based Health Information Technology: Access Among Latinos Varies by Subgroup Affiliation (NHIS 2015-2016) https://www.jmir.org/2019/4/e10389/
- Zhang2017: Comparison of Health Information Technology Use Between American Adults With and Without Chronic Health Conditions (NHIS 2012) https://www.jmir.org/2017/10/e335/

Accessing NHIS data

- Data briefs https://www.cdc.gov/nchs/products/databriefs.htm
- QuickStats https://www.cdc.gov/nchs/nhis/nhis_quickstats.htm
- Early Release Measures https://www.cdc.gov/nchs/nhis/releases/released201905.htm
- Data to download https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm
- 2018 data: https://www.cdc.gov/nchs/nhis/nhis_2018_data_release.htm

Considerations

- Cross-sectional
- Self-reported data
- Data on health insurance, medical conditions, health care utilization, household composition
- Large sample
- Nationally representative
- Household data, parent-child dyads
- Can combine across years
- Linkages with other datasets

Current Population Survey (CPS)

- Sponsored jointly by the Census Bureau and Bureau of Labor Statistics
  - https://www.census.gov/programs-surveys/cps.html
  - https://www.bls.gov/cps/
- Primary source of labor force statistics for the US
  - Labor force, employment, unemployment, persons not in the work force, hours of work, earnings, demographics
- Probability sample of 60,000 households
  - Designed to produce national and state estimates of labor force characteristics in civilian non-institutionalized population 16+ years
- Supplemental surveys: food security, tobacco use, fertility, computers & Internet use
  - https://www.census.gov/programs-surveys/cps/about/supplemental-surveys.html

https://www.census.gov/programs-surveys/cps/technical-documentation/methodology.html

Questions of interest

- Use of desktop computer, laptop, tablet, wearable devices; wifi, broadband, dial-up
- Internet use at home, work, school, café, on public transit, library/other community location
- Use email, texting/instant messaging, social media, watch videos, radio/stream music
- Connected household devices (e.g., thermostat)
- Communicate with healthcare provider, access medical records
- Concerns about online privacy, cyberbullying
- Reasons not online at home

https://www2.census.gov/programs-surveys/cps/techdocs/cpsnov17.pdf

Example Publications

- Wright2009: Prescription for trouble: Medicare Part D and patterns of computer and internet access among the elderly https://www.ncbi.nlm.nih.gov/pubmed/19333841
- Ryan2018: Computer and Internet Use in the United States: 2016 https://www.census.gov/content/dam/Census/library/publications/2018/acs/ACS-39.pdf (report)
- National Telecommunications and Information Association, 2014: Exploring the Digital Nation: Embracing the Mobile Internet https://www.ntia.doc.gov/files/ntia/publications/exploring_the_digital_nation_embracing_the_mobile_internet_10162014.pdf

Accessing CPS data

- Data available via Census and BLS websites
- https://www.bls.gov/cps/data.htm
- Table creator https://www.census.gov/cps/data/cpstablecreator.html
  - In progress of migrating to new tool
- DataFerrett https://dataferrett.census.gov/?#
  - Data analysis and extraction tool to customize federal, state, and local data to suit your requirements
  - Migration to new tool in progress https://data.census.gov/mdat/#/
- Micro data https://thedataweb.rm.census.gov/ftp/cps_ftp.html?#

Considerations

- Large sample size
- Focus on employment and demographics
- Limited health data

FLASHE

Family Life, Activity, Sun, Health, and Eating Study

- https://cancercontrol.cancer.gov/brp/hbrb/flashe.html
- Survey data of psychosocial, generational (parent-adolescent), and environmental correlates of cancer-preventive behaviors
- Cross-sectional Internet-based survey in 2014
- Adolescent and parent/caregiver completed 2 web surveys (n=1479 dyads)
- Balanced sampling: sample to match US households with 1+ 12-17yo as closely as possible on gender, Census division, household income, household size, and race/ethnicity

https://cancercontrol.cancer.gov/brp/hbrb/docs/FLASHE_Methods_Report.pdf

Data collection

- Teen diet survey
  - Attitudes & opinions
  - What you eat and drink
  - Food away from home
  - Food in your home
  - Family meals
  - Your preferences
  - Your parents
- Teen physical activity survey
  - Physical activity
  - Your home and neighborhood
  - Using electronic devices
  - Time spent in the sun and indoor tanning
  - Tobacco use
  - Sleep
  - Goals in life
  - Your parents

Parent/caregiver surveys – similar topics/questions

https://cancercontrol.cancer.gov/brp/hbrb/flashe-files.aspx

Questions of Interest

- Teen & Parent
  - Hours/day use computer, phone, TV, gaming devices, electronic reader in school and outside of school
  - Why would limit use of electronic devices, confidence in ability to limit use
- Parent
  - Teen's use of computer, phone, TV, gaming devices, electronic reader
  - Parenting rules about screen time

https://cancercontrol.cancer.gov/brp/hbrb/docs/Teen_PA_PUF_Instrument.pdf
https://cancercontrol.cancer.gov/brp/hbrb/docs/Parent_PA_PUF_Instrument.pdf

Example studies

- Zhang2019: parenting styles and adolescents' energy balance-related behaviors https://www.ncbi.nlm.nih.gov/pubmed/31299191
- Joyal-Desmarais2019: Interpersonal effects of parents and adolescents on each other's health behaviours https://www.ncbi.nlm.nih.gov/pubmed/30618314
- Yang2019: The Effect of Screen Viewing Duration and Self-Efficacy in Limiting Screen Viewing on Loneliness in Adolescent-Parent Dyads https://www.ncbi.nlm.nih.gov/pubmed/31108323
- Rice2019: Interactions among perceived norms and attitudes about health-related behaviors in U.S. adolescents https://www.ncbi.nlm.nih.gov/pubmed/30762406

Accessing FLASHE data

- https://cancercontrol.cancer.gov/brp/hbrb/flashe-files.aspx
- SAS code and webinar on how to conduct dyadic analyses https://cancercontrol.cancer.gov/brp/hbrb/flashe-dyadic-analysis.aspx

Considerations

- Parent/caregiver-adolescent dyads
- Cross-sectional
- Data self-reported
  - Except for adolescent accelerometer data
- Neat data on cancer risk factors: behaviors, motivations, opinions, parenting rules
- Non-probability sample from volunteer online consumer opinion panel – similar characteristics to US households with adolescents 12-17 years old
  - Parent/caregivers mostly female, highly educated

Youth Risk Behavior Survey (YRBS)

- https://www.cdc.gov/healthyyouth/data/yrbs/index.htm
- National school-based survey conducted by CDC and state, territorial, tribal, and local surveys conducted by state, territorial, and local education and health agencies and tribal governments
- Conducted every 2 years – data available for 1991-2017
- Datasets available: national, by state, large school districts
- Representative of 9th-12th grade students in public and private schools in the US

Data Collection

- Monitors 6 categories of health-related behaviors
- Survey instruments https://www.cdc.gov/healthyyouth/data/yrbs/questionnaires.htm
- Questions of interest
  - Texting or emailing while driving
  - Video games/texting/social media use
  - Cyberbullying

Standard and National High School Questionnaires
- Demographics
- Height and Weight
- Unintentional Injuries and Violence
- Tobacco Use
- Electronic Vapor Product Use
- Alcohol and Other Drug Use
- Sexual Behaviors
- Weight Management
- Dietary Behaviors
- Physical Inactivity
- HIV
- Other Topics

Middle School Questionnaire
- Demographics
- Height and Weight
- Unintentional Injuries and Violence
- Tobacco Use
- Electronic Vapor Product Use
- Alcohol and Other Drug Use
- Sexual Behaviors
- Weight Management
- Dietary Behaviors
- Physical Inactivity
- Other Topics

Example studies

- Rostad2018: Association Among Television and Computer/Video Game Use, Victimization, and Suicide Risk Among U.S. High School Students (YRBS 2015) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6119526/
- Alhajji2019: Cyberbullying, Mental Health, and Violence in Adolescents and Associations With Sex and Race (YRBS 2015) https://www.ncbi.nlm.nih.gov/pubmed/31431904
- Baiden2019: The association between excessive screen-time behaviors and insufficient sleep among adolescents (YRBS 2017) https://www.ncbi.nlm.nih.gov/pubmed/31629305

Accessing Data

- Online data analysis tool
  https://nccd.cdc.gov/youthonline/App/Default.aspx
  - All US or by state, sex, race/ethnicity, sexual orientation, grade
- Download data (1991-2017)
  https://www.cdc.gov/healthyyouth/data/yrbs/data.htm
- Results https://www.cdc.gov/healthyyouth/data/yrbs/results.htm
- Infographics https://www.cdc.gov/healthyyouth/data/yrbs/toolkit.htm
- Reports of 2017 YRBS
  https://www.cdc.gov/healthyyouth/data/yrbs/pdf/2017/ss6708.pdf

Played Video Or Computer Games Or Used A Computer For 3 Or More Hours Per Day
(counting time spent on things such as Xbox, PlayStation, an iPad or other tablet, a smartphone, texting, YouTube, Instagram, Facebook, or other social media, for something that was not school work, on an average school day)
United States, High School Youth Risk Behavior Survey, 2017

Find out if there is a statistical difference between two grades. Select two and activate 'Compare Two'. COMPARE TWO »

Year \ Grade	Total	9th	10th	11th	12th
2017	43.0 (41.1–44.9) 13,839†	45.0 (41.7–48.2) 3,636	45.1 (42.3–47.8) 3,480	42.3 (38.5–46.1) 3,430	39.2 (36.7–41.8) 3,165

National Survey of Children's Health

- http://childhealthdata.org/learn-about-the-nsch/NSCH
- https://mchb.hrsa.gov/data/national-surveys
- Designed to produce national and state-level data on the physical & emotional health of US children 0-17 years
  - Related factors: medical home, family interactions, parental health, school experiences, safe neighborhoods
- Non-institutionalized children 0-17 years sampled from all 50 states and Washington DC
- Adult who knows the child best (usually a parent) completes online or paper survey about the child
  - 2003, 2007, 2011-2012: 91K-102K nationally, 1,800-2,200 per state
  - 2016: 50,212 nationally, 638-1,351 per state
  - 2017: 21,599 nationally, 343-454 per state

http://childhealthdata.org/learn-about-the-nsch/NSCH
http://childhealthdata.org/learn-about-the-nsch/survey-revisions

Data collection

- SECTION 1: Initial Screener
- SECTION 2: Survey Questions
  - A. This Child's Health
  - B. This Child as an Infant
  - C. Health Care Services
  - D. Experience with This Child's Health Care Providers
  - E. This Child's Health Insurance Coverage
  - F. Providing for This Child's Health
  - G. This Child's Learning (0-5 years)
  - G. This Child's Schooling and Activities (6-17 years)
  - H. About You and This Child
  - I. About Your Family and Household
  - J. About You
  - K. Household Information

Questions of interest

- On an average weekday, about how much time does this child usually spend:
- … in front of a TV watching TV programs, videos, or playing video games?
- …with computers, cell phones, handheld video games, and other electronic devices, doing things other than schoolwork?

http://childhealthdata.org/learn-about-the-nsch/topics_questions/2017-nsch-guide-to-topics-and-questions
http://childhealthdata.org/learn-about-the-nsch/survey-instruments

Accessing data

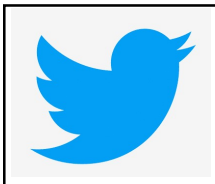
Indicator 6.11: On an average weekday, about how much time does this child usually spend with computers, cell phones, handheld video games, and other electronic devices, doing things other than schoolwork?

	Do not use electronic devices	Use electronic devices less than 1 hour per day	Use electronic devices 1 hour per day	Use electronic devices 2-3 hours per day	Use electronic devices 4 hours or more per day	Total %
%	17.6	19.6	21.1	30.9	10.9	100.0
C.I.	16.4 - 18.7	18.5 - 20.8	19.9 - 22.3	29.5 - 32.2	10.1 - 11.7	
Sample Count	3,235	4,199	4,592	6,833	2,543	
Pop. Est.	12,746,370	14,228,158	15,319,684	22,414,029	7,901,437	

- Interactive data query
  - http://childhealthdata.org/browse/survey
  - Survey year, US vs state-level
  - Sample N, weighted N, weighted % (95% CI) provided
- Public-use datasets
  - https://www.census.gov/data/datasets/2017/demo/nsch/nsch2017.html
  - Includes variable list, methodology report, analytic guides
  - Note, major changes to survey administration in 2016; pre-2016 vs 2016+ surveys not comparable

Example studies

- Healy2020: Physical Activity, Screen Time, and Sleep Duration Among Youth With Chronic Health Conditions in the US (NSCH 2016) https://www.ncbi.nlm.nih.gov/pubmed/32233773
- Lo2015: TV in bedroom and weekday screen time in kids with ADD/ADHD (NSCH 2007) https://www.sciencedirect.com/science/article/pii/S221133551400014X
- Sisson2009: Profiles of sedentary behavior in children and adolescents (NSCH 2001-2006) https://www.ncbi.nlm.nih.gov/pubmed/19922052



Twitter

- Public tweets
- Examine what people are saying about a particular health-related topic
- Results generalizable to people who tweet about that topic using keywords/hashtags sampled

Check out our previous webinar: "Methods for Capturing and Examining Social Media Data for Health Research" https://www.youtube.com/watch?v=B-wMT151NrY

Accessing Twitter data

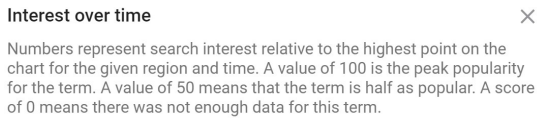
- Data collection tools
  - NCapture, NVivo add-on for Chrome https://www.qsrinternational.com/nvivo/support-overview/faqs/what-is-ncapture
  - R https://rtweet.info/
  - For-free services (e.g., Gnip)
- Sample posts by hashtag, search term, or user
- Considerations
  - Sampling methods
  - Have to authorize through your social media account
  - Protect user privacy

Example studies

- Kim2020: At the speed of Juul: measuring the Twitter conversation related to ENDS and Juul across space and time (2017-2018) https://www.ncbi.nlm.nih.gov/pubmed/32198278
- Griffis2020: Using Social Media to Track Geographic Variability in Language About Diabetes: Analysis of Diabetes-Related Tweets Across the US https://diabetes.jmir.org/2020/1/e14431/
- Merrill2020: Posting Post-Blackout: A Qualitative Examination of the Positive and Negative Valence of Tweets Posted After "Blackout" Drinking https://pubmed.ncbi.nlm.nih.gov/31986999/

Google Trends

- https://trends.google.com/trends
- Google search patterns January 2004 to present
- Search by location worldwide, date, categories, or type of Google search
- RSV = relative search volume
- Accessing data
  - Download data as CSV file
  - Create graphs/charts



Considerations

- Represents Google searches on a topic during time period in geographic area
  - Searches, not people
  - Doesn't include other search engines or other ways to looking for (health-related) information online
- Measure of public interest in topic
- Selecting search terms used to describe concept
- Long-term or seasonal trends
- Compare across geographic region

Example studies

- Niforatos2019: public interest in gun control in the USA https://www.ncbi.nlm.nih.gov/pubmed/30992330
- Caputi2018: Google Searches for "Cheap Cigarettes" Spike at Tax Increases: Evidence from an Algorithm to Detect Spikes in Time Series Data https://academic.oup.com/ntr/article/20/6/779/3884451
- Seth2018: Association of socioeconomic and geographic factors with Google trends for tanning and sunscreen https://www.ncbi.nlm.nih.gov/pubmed/28902024
- Madden2017: The seasonal periodicity of health contemplations about exercise and weight loss: ecological correlational study https://publichealth.jmir.org/2017/4/e92/
- Jo2015: US consumer interest in non-cigarette tobacco products spikes around the 2009 federal tobacco tax increase https://www.ncbi.nlm.nih.gov/pubmed/24500270

Overview of existing datasets

- Adults: HINTS, Pew, NHIS, CPS
- Children/adolescents: FLASHE, YRBS, NSCH
- Posts/searches: Twitter, Google trends

Interested in learning more about research using large population-based datasets?

- AH6015: Use of Large Population-Based Datasets for Health Promotion
  - Fall 2020 Tuesdays 1:30-4pm
- MS and PhD programs in Health Promotion Sciences
  - https://healthpromotionsciences.uconn.edu/

UConn Center for mHealth and Social Media Annual Conference

- "Building an Evidence Base for Commercially Available Technology" – VIRTUAL this year – 14-15 May 2020
- More info and registration https://mhealth.inchip.uconn.edu/events/
- Agenda, speakers, workshops https://mhealth.inchip.uconn.edu/uconndigital2020/
- Abstract submissions due Fri 4/10
  - 1-minute video "poster"