



Database Research: Best Practices

## Databases for surgical health services research: National Health and Nutrition Examination Survey



### Introduction

For more than half a century the National Health and Nutrition Examination Survey (NHANES) has been a source population for tens of thousands of peer-reviewed studies.<sup>1</sup> Historically these have been *cross-sectional studies* featuring mostly nonsurgical topics; however, some have covered topics in urology, otolaryngology, bariatric surgery, surgical oncology, and other surgical specialties.<sup>2</sup> In the last decade, at least six *cohort studies* using NHANES and featuring topics in transplant surgery have been published in top-tier journals.<sup>3–8</sup> NHANES is a potentially valuable resource for surgeons and NHANES databases may be only a click away!

### What is NHANES?

NHANES is a series of nationally representative surveys conducted by the United States' National Center for Health Statistics (NCHS) between 1959 and 2018.<sup>9</sup> An NHANES data file may include up to 3000 fields of individual-level medical information obtained from interviews, physical examination, and imaging and laboratory tests at the time of participant enrollment (Table 1).

### What questions can NHANES data help surgeons answer?

By design, NHANES is ideally suited for survey-type studies that describe the *prevalence* of the health and nutritional status of children and adults in the US, such as prevalence of overweight and obesity.<sup>10</sup> Using NHANES-derived estimates of the prevalence of obesity, surgical researchers have projected that 14.7% of obese adults in the US (approximately 140 million) should be considered for bariatric surgery.<sup>11</sup> This extraordinary inference may influence health policy and planning, as well as medical education planning, since it points to a very high demand for bariatric surgeons. Yet obesity is merely 1 of 3000 fields in the NHANES database; as such, it is plausible that hundreds to thousands of other topics of interest to the surgeon may be similarly investigated using NHANES.

NHANES is also ideally suited for cross-sectional studies that assess the *association* between health characteristics and clinical outcomes at a given time. Recently, ENT-surgeons at Johns Hopkins University used NHANES data to assess the association between tobacco smoke exposure and Eustachian tube disorders in US children and adolescents.<sup>12</sup> The essence of such an *association study* is the exposure of interest (tobacco smoking in this example),

the outcome of interest (Eustachian tube disorders), and the potential confounders of the association under study (age, sex, body mass index, education level, ethnicity; having a cold, sinus problem, or ear ache during the last 24 hours). Because tobacco smoking and Eustachian tube disorders are merely 2 of 3000 fields in the NHANES database, it is likely that hundreds to thousands of other topics of interest to the surgeon may be similarly investigated using NHANES.

Fairly recent NCHS innovations have enabled researchers to go beyond cross-sectional design to implement *longitudinal* cohort analyses with NHANES data (Fig. 1). Survey participant data (captured from a cross-section of time) have been linked using social security numbers to CMS Medicare enrollment and claims files (for records on subsequent initiation of dialysis or receipt of a kidney transplant) and to the national death index (to ascertain death). Thus, researchers can observe or follow NHANES participants from the date of survey enrollment to the end of observation, when the survey data were linked to external databases.

The essence of a cohort study is the date of enrollment (in this case the date when the NHANES survey was administered), the date at the end of the observation period (in this case the date when the NHANES database was linked to the CMS and NDI databases to ascertain ESRD and death), participant status at the end of observation (developed ESRD, died, or did not develop either outcome), and the baseline confounders of interest (thousands of variables to choose from). Kidney transplant researchers have leveraged NHANES data linkages and published several studies that have influenced clinical science, practice, and guidelines in the US and across the world.<sup>13</sup>

### How may surgeons access NHANES data?

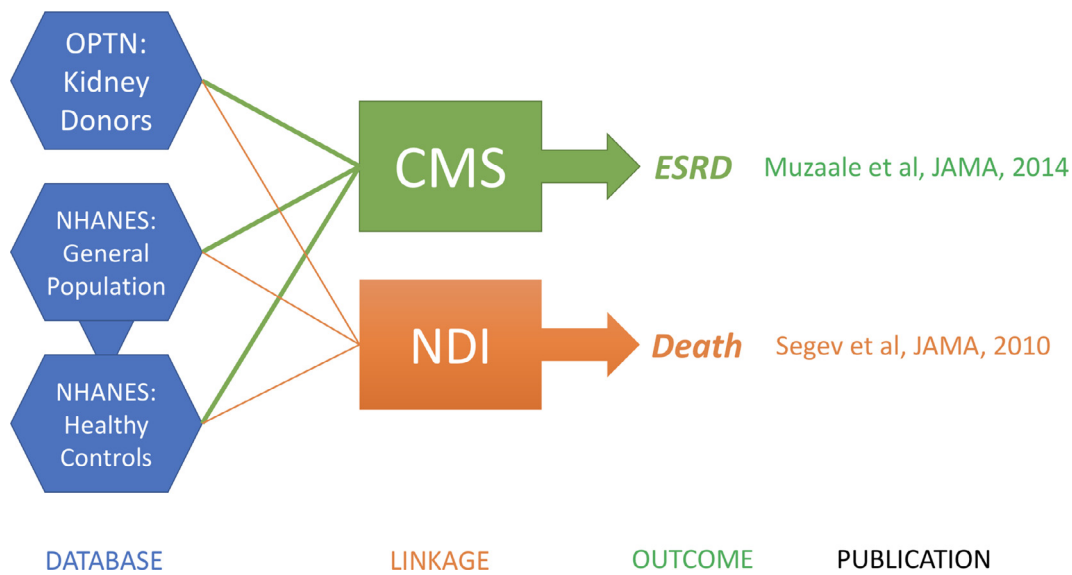
*Public-use data files* are free and downloadable from the NCHS website and users may have access to datasets, documentation, and questionnaires from the NHANES survey and other NCHS data collection systems.<sup>14</sup>

*NHANES data linkages* are available for researchers who wish to longitudinally examine the factors that influence disability, chronic disease, health care utilization, morbidity, and mortality. NCHS is linking NHANES with several databases including air-monitoring data from the Environmental Protection Agency, death certificates from the National Death Index, Medicare enrollment and claims data from the Centers for Medicare & Medicaid Services, and retirement, survivor, and disability insurance and Supplemental

**Table 1**  
The National Health and Nutrition Examination Surveys: 1959–2018.

Data file	Survey	Focus
NHES I	1959–1962	Chronic disease in adults aged 18–79
NHES II	1963–1965	Growth and development in children aged 6–11
NHES III	1966–1970	Growth and development of adolescents aged 12–17
NHANES I	1971–1974	Added a nutrition component; participants aged 1–74
NHANES II	1976–1980	Participants aged 6 months to 74 years
Hispanic HANES	1982–1984	Mexican-American, Cuban-American, Puerto Rican
NHANES III	1988–1994	Oversampled many groups. Ages 2 months and over
Continuous NHANES <sup>a</sup>	1999–2018	Search NCHS website for variables <sup>17</sup>

<sup>a</sup> Conducted by the NCHS every year since 1999.



**Fig. 1.** Illustrates how the National Health and Nutrition Examination Survey (NHANES) database has been used in kidney transplant research. The Organ Procurement and Transplant Network (OPTN) was used as the source population for live kidney donors (i.e., study population) and NHANES was used as the source population for matched healthy controls (i.e., control population). The OPTN and NHANES databases were linked to the Centers for Medicare and Medicaid Services (CMS) and to the National Death Index (NDI) databases by federal employees at the respective agencies. These data linkages permitted the study of nephrectomy-attributable perioperative and long-term risks following live kidney donation.

Security Income benefit data from the Social Security Administration.<sup>15</sup>

The *National Death Index (NDI)* is a national database of death record information. Using records submitted by state vital statistics offices, NDI may aid investigators with their mortality-related research.<sup>3–5</sup> Death records are added to the NDI file annually, approximately 12 months after the end of each calendar year.

The *Research Data Center (RDC)* provides private-access NHANES data files for researchers comfortable using SAS, Stata, SPSS or similar statistical packages. Researchers must submit a proposal outlining their need for data that are more sensitive than the otherwise free, downloadable public-use files. Once the proposal is approved, RDC works with researchers to create a data file specific to the research question. Although the RDC cannot send such private-access datasets offsite, it offers several options for accessing the file. The RDC has been an invaluable resource to kidney transplant researchers over the last five years.<sup>7,8,16</sup>

## Summary

NHANES is a series of nationally representative surveys conducted by the NCHS between 1959 and 2018. NHANES data have up to 3000 fields of individual-level medical information obtained from interviews, physical examination, and imaging and laboratory tests at the time of participant enrollment. It is one of the most accessible databases for studying topics of relevance to the surgeon. Major surgical research contributions using NHANES have

demanded nothing more than access to the NCHS website, access to commercially available statistical software, and an imaginative idea inspired from the trenches of clinical work.

Abimereki D. Muzaale, MD, MPH\*

Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD

\*Corresponding author.

E-mail address: [amuzaal1@jhmi.edu](mailto:amuzaal1@jhmi.edu)

Accepted 13 February 2018

## References

1. Pubmed best match for 'NHANES'. <https://www.ncbi.nlm.nih.gov/pubmed/?term=NHANES>; 2018 Accessed February 1.
2. Pubmed best match for 'NHANES Surgery'. <https://www.ncbi.nlm.nih.gov/pubmed/?term=NHANES+surgery>; 2018 Accessed February 1.
3. Ibrahim HN, Foley R, Tan L, Rogers T, Bailey RF, Guo H, et al. Long-term consequences of kidney donation. *N Engl J Med.* 2009;360:459–469.
4. Segev DL, Muzaale AD, Caffo BS, Mehta SH, Singer AL, Taranto SE, et al. Perioperative mortality and long-term survival following live kidney donation. *JAMA.* 2010;303:959–966.
5. Berger JC, Muzaale AD, James N, Hoque M, Wang JM, Montgomery RA, et al. Living kidney donors ages 70 and older: recipient and donor outcomes. *Clin J Am Soc Nephrol.* 2011;6:2887–2893.
6. Muzaale AD, Dagher NN, Montgomery RA, Taranto SE, McBride MA, Segev DL. Estimates of early death, acute liver failure, and long-term mortality among live liver donors. *Gastroenterology.* 2012;142:273–280.

7. Muzaale AD, Massie AB, Wang MC, Montgomery RA, McBride MA, Wainright JL, et al. Risk of end-stage renal disease following live kidney donation. *JAMA*. 2014;311:579–586.
8. Grams ME, Sang Y, Levey AS, Matsushita K, Ballew S, Chang AR, et al. Kidney-failure risk projection for the living kidney-donor candidate. *N Engl J Med*. 2016;374:411–421.
9. National Health and Nutrition Examination Survey. <https://wwwn.cdc.gov/nchs/nhanes/Default.aspx>; 2018 Accessed February 1.
10. Ogden CL, Carroll MD, Curtin LR, McDowell MA, Tabak CJ, Flegal KM. Prevalence of overweight and obesity in the United States, 1999–2004. *JAMA*. 2006;295:1549–1555.
11. Stevens J, Oakkar EE, Cui Z, Cai J, Truesdale KP. US adults recommended for weight reduction by 1998 and 2013 obesity guidelines, NHANES 2007–2012. *Obesity (Silver Spring)*. 2015;23:527–531.
12. Patel MA, Mener DJ, Garcia-Esquinas E, Navas-Acien A, Agrawal Y, Lin SY. Tobacco smoke exposure and Eustachian tube disorders in US children and adolescents. *PLoS One*. 2016;11.
13. Steiner RW. A very different paradigm for living kidney donor risk. *Am J Transplant*. 2017;17:1701–1702.
14. NCHS data linkage. <https://www.cdc.gov/nchs/data-linkage/new-linkage.htm>; 2018 Accessed February 1.
15. NCHS Resources for researchers. [https://www.cdc.gov/nchs/nchs\\_for\\_you/researchers.htm](https://www.cdc.gov/nchs/nchs_for_you/researchers.htm); 2018 Accessed February 1.
16. Muzaale AD, Althoff KN, Sperati CJ, Abraham AG, Kucirka LM, Massie AB, et al. Risk of end-stage renal disease in HIV-positive potential live kidney donors. *Am J Transplant*. 2017;17:1823–1832.
17. Search variables in continuous NHANES. <https://wwwn.cdc.gov/nchs/nhanes/search/default.aspx>; 2018 Accessed February 1.