TITLE: Cardiometabolic statistics for US adults: an open source platform for exploratory analysis

AUTHORS (unordered): Paul Muntner, Ligong Chen, Byron Jaeger, <OTHERS FROM JHS HWG WRITING RETREAT>

INTRODUCTION

* Describe the NHANES, quantify its impact:
  + How many papers use NHANES data?
  + how many healthcare policies/guidelines in the US are based on NHANES data?
* Identify problem: NHANES data are publicly available but are difficult to analyze. Challenges include data ascertainment, variable derivation, inclusion and exclusion of participants, carrying out valid statistical analysis accounting for the complex survey design of NHANES.
* Proposed solution: We develop a free online platform that allows users to obtain visual and tabular summaries of the NHANES data using validated statistical techniques.

METHODS

* The NHANES data and protocols for data collection
* NHANES data used in the current analysis:
  + Which survey participants were included (**Figure S1**)
  + NHANES data we included and variables derived from these data (**Table S1**)
* What can the app do?
  + Generate tabular or visual summaries of NHANES data that can be saved by the user
  + Pool results over contiguous NHANES cycles or stratify results by cycle
    - If stratifying by cycle, user can adjust for age differences using direct standardization
  + Subset the data to be summarized, e.g., by including or excluding adults who report taking antihypertensive medications.
  + Stratify results across subgroups, e.g., compute the mean systolic blood pressure among men and women, separately.
* Development and validation of the app
  + We used R version 4.1.2 to develop a web application using the “Shiny” R package
  + We disseminate the app publicly online
  + We disseminate code and data from the app in an open-source R Package
  + Tests for validity of the app included replication of summary statistics from
    - From Muntner et al, JAMA 2020
    - From <https://www.cdc.gov/nchs/data/databriefs/db289.pdf>
    - From Carey et al

RESULTS

Note: all results do not include women who were pregnant

* Prevalence of hypertension **in the overall sample**
  + By calendar period (**Figure 1**)
  + By age group, sex, and race-ethnicity in 2017-2020 (**Figure S2**)
* Antihypertensive medication **among those with hypertension**
  + Prevalence of antihypertensive medication use by calendar period (**Figure 2**)
  + Number of antihypertensive medication classes by calendar period (**Figure S4**)
* Prevalence of different antihypertensive medications being taken **among those with hypertension**
  + Identify 3 or 4 specific medications
* Prevalence of resistant hypertension among those with hypertension
  + By calendar period (**Figure 3**)
  + By age group, sex, and race-ethnicity in 2017-2020 (**Figure S5**)
* Blood pressure control **among those with hypertension**
  + Prevalence of blood pressure control (JNC7 defn) by calendar period (same as in JAMA; **Figure 4**)
  + Prevalence of blood pressure control (JNC7 defn) by age group, sex, and race-ethnicity in 2017-2020 (**Figure S6**)
* Mean SBP and DBP
  + By calendar year (**Figure S7**)
  + By age group, sex, race/ethnicity (**Figure S8**)
* A supplemental “tutorial video” where the first author shows how to make all of the main figures using the shiny app?

DISCUSSION

* First paragraph: summary.
  + We made an application that allows anyone to obtain valid summaries of US population statistics for cardiometabolic outcomes.
  + We tested the application by replicating prior work with it
  + Our results show that one can replicate prior NHANES results and extend them
* Second paragraph: extending this app
  + We have presented BP outcomes but have designed the app to be modular and we plan to add several other cardiometabolic outcomes in future updates.
  + We plan on incorporating the 2021-2022 NHANES cycle as soon as possible, following its public release
  + Others can develop extensions to this application by collaborating with us on GitHub
* Third paragraph: Who can benefit from this app
  + In research settings, many population scientists may wish to engage with NHANES data, but do not due to the difficulty of and time required to retrieve the data, derive the relevant variables, and account for the complex survey design of NHANES.
  + In educational settings, many teachers in public health may wish to show recent trends in health outcomes, but do not due to the same obstacles above.
  + Investigators writing grant submissions may wish to cite recent statistics regarding US adults, but it is difficult to identify prior papers providing the specific data they want to report.

MAIN TABLES

**Table 1**: characteristics of NHANES participants included in the data for the web application.

MAIN FIGURES

TO BE FILLED IN LATER

SUPPLEMENTAL TABLES

**Table S1**: Variables used in the current analysis

|  |  |  |
| --- | --- | --- |
| Variable | Source | Description |
| Demographics | | |
| Age, years | DEMO | Age in years of the survey participant at the time of screening. Individuals 80 years or older are coded at 80 years of age. |
| Age category, years | Derived | Five categories are derived:  18 to 44 years  45 to 64 years  65 to 74 years  75+ years |
| Race | DEMO | Non-Hispanic White  Non-Hispanic Black  Non-Hispanic Asian  Hispanic  Other |
| Pregnant | DEMO | Yes  No |
| Gender | DEMO | Female  Male |
| Blood Pressure | | |
| SBP, mm Hg | BPX | Three consecutive BP measurements (systolic and diastolic) were taken 60 seconds apart |
| DBP, mm Hg | BPX |
| BP category | Derived | SBP <120 and DBP <80 mm Hg  SBP of 120 to <130 and DBP <80 mm Hg  SBP of 130 to <140 or DBP 80 to <90 mm Hg  SBP of 140 to <160 or DBP 90 to <100 mm Hg  SBP 160+ or DBP 100+ mm Hg |
| BP control |  |  |
| JNC7 definition | Derived | SBP < 140 mm Hg and DBP < 90 mm Hg |
| 2017 ACC/AHA definition | Derived | SBP < 130 mm Hg and DBP < 80 mm Hg |
| Uncontrolled BP |  |  |
| JNC7 definition | Derived | SBP >= 140 mm Hg or DBP >= 90 mm Hg |
| 2017 ACC/AHA definition | Derived | SBP >= 130 mm Hg or DBP >= 80 mm Hg |

BP = blood pressure, SBP = systolic blood pressure, DBP = diastolic blood pressure

SUPPLEMENTAL FIGURES

**Figure S1**: cascade of inclusions, use a similar design as the inclusion/exclusion cascade in the JAMA paper

Figures S2 and above to be filled in later