

All official NCHS BSC documents are posted on the BSC website
(https://www.cdc.gov/nchs/about/bsc/bsc_meetings.htm)

Department of Health and Human Services (HHS)

**Board of Scientific Counselors
National Center for Health Statistics
Centers for Disease Control and Prevention
January 27, 2021**

Meeting Summary

The Board of Scientific Counselors (BSC) convened via Zoom on January 27, 2021. The virtual meeting was open to the public (via Zoom).

Board Members Present

Linette T. Scott, M.D., M.P.H., Chair, BSC
Kennon R. Copeland, Ph.D.
Robert M. Hauser, Ph.D.
Scott H. Holan, Ph.D.
Helen G. Levy, Ph.D.
R. John Lumpkin, M.D., M.P.H. [joined ~12:15pm ET]
Sally C. Morton, Ph.D.
Kristen M. Olson, Ph.D.
Andrey Peytchev, Ph.D.
Gretchen Van Wye, Ph.D., M.A.

CDC/NCHS Panelists

Brian Moyer, Ph.D., Director, National Center for Health Statistics (NCHS)
Sayeedha Uddin, M.D., M.P.H., Designated Federal Officer, NCHS, BSC
Irma Arispe
Stephen Blumberg
Amy Branum
James Craver
Carol DeFrances
Renee Gindi
Travis Hoppe
David Huang
Florence Lee
Gwen Mustaf
Tina Norris
Cynthia Ogden
Lucila Ohno-Machado
Jennifer Parker
Ryne Paulose
Matthew Snipp
Paul Sutton
Brian Ward

Other Attendees

Naman Ahluwalia	Yulei He	Van Parsons
Lara Akinbami	Holly Hedegaard	Omar Pedraza
Johanna Alfier	Elizabeth Heitz	Sue Pedrazzani
Josephine Alford	Kevin Heslin	Zachary Peters
Nick Ansai	Emma Heth	Kellina Phan
Yutaka Aoki	Rebecca Hines	Steve Pierson
Jonathan Aram	MaryJo Hoeksema	John Pleis
Jill Ashman	Jacquie Hogan	Paul Pulliam
Brenda Baker	Robert Hood-Cree	Cynthia Reuben
Bryan Bassig	John Hough	Minsun Riddles
Joseph Baweja	Rebecca Hu	Richard Riley
Greg Binzer	Lori Hunter	Darryl Roberts
Peter Boersma	Deborah Ingram	Charles Rothwell
Jonaki Bose	Katherine Irimata	Neil Russell
Jim Brittain	Matt Jans	Katie Russo
Deb Brody	Christine Jones	Asel Ryskulova
Lisa Broitman	Jessly Joy	Neda Sarafrazi
Sherry Brown Scoggins	Ed Kilbourne	Paul Scanlon
Verita Buie	Christine Kim	Paul Schroeder
Eileen Call	Jieun Kim	Steven Schwartz
Te-Ching Chen	Richard Klein	Bobbie Shimizu
Yao Cheng	Karen Knight	Lisa Simpson
Chanda Chay	Ellen Kramarow	Tammy Stewart-Prather
Kristen Cibelli	Nataliya Kravets	Bryan Stierman
Jodi Cisewski	Melissa Kresin	Renee Storandt
Jay Clark	Tom Krenzske	Yu Sun
Christine Cox	Deanna Kruszon-Moran	Makram Talih
Lauren Creamer	Michael Lawrence	Rashmi Tandon
Nicole Cummings	Xianfen Li	Chally Tate
Lucie Dalzell	Susan Lukacs	Ana Terry
Barnali Das	Aaron Maitland	Tim Tilet
Danielle Davis	Donald Malec	Daouda Traore
Orlando Davy	Bradley Malin	Allan Uribe
Debbie Dean	Carroll Margaret	Anjel Vahratian
Jerry Del Rosso	Crescent Martin	Lisa Wagner
Hua Di	Juliana McAllister	Edwina Wambogo
Michele Dillon	Karen Melia	Cha-Yih Wang
Natalie Dupree	Justin Mezetin	Antonia Warren
Morgan Earp	Donna Miller	Julie Weeks
Nazik Elgaddal	Kristen Miller	Kristen Welker-Hood
Steven Fink	Lisa Mire	Jean Williams
Sheila Franco	Suruchi Mishra	Sonja Williams
Alicia Frasier	Leyla Mohadjer	Ashley Woodall
Chris Freedman	Mary Moien	Shannon Woodis
Cheryl Fryar	Chris Moriarity	David Woodwell
Doreen Gidali	Kiana Morris	Evangeline Woody
Cordell Golden	Kelly Myrick	Alana Yick
Jessica Graber	Zakia Nelson	Carla Zelaya
Lauren Gross	Colleen Nugent	Keith Zevallos

Qiuping Gu	Tatiana Nwankwo	Call-in Participants (n=14)
Brady Hamilton	Kathleen O'Connor	
Althelia Harris	Damon Ogburn	
LaJeana Hawkins	Vera Osidach	
Qiming He	Zoe Padgett	
Tina Baker (Captioner)	Greg Richards (Meeting Host, Rose Li & Associates)	Dana Gleit (Minutes, Rose Li & Associates)

List of Abbreviations

ACS	American Community Survey
API	Application Programming Interface
BSC	Board of Scientific Counselors
CDC	Centers for Disease Control and Prevention
CHC	Community Health Center
COVID-19	Coronavirus Disease 2019
DAE	Division of Analysis and Epidemiology
DHANES	Division of Health and Nutrition Examination Surveys
DHCS	Division of Health Care Statistics
DMI	Data Modernization Initiative
EHR	Electronic Health Record
FY	Fiscal Year
HHS	U.S. Department of Health and Human Services
HTML	Hypertext Markup Language
NAMCS	National Ambulatory Medical Care Survey
NCHS	National Center for Health Statistics
NHANES	National Health and Nutrition Examination Survey
NHIS	National Health Interview Survey
NVSS	National Vital Statistics System
PDF	Portable Document Format
PSU	Primary Sampling Units
RANDS	Research and Development Survey
RDC	Research Data Center

Action Steps

- The BSC voted unanimously to endorse the BSC National Hospital Ambulatory Medical Care Survey (NAMCS) Workgroup report and submit its recommendations based on the NAMCS Workgroup’s responses to the questions posed by the NAMCS program team.
- Dr. Moyer encouraged BSC members to contribute feedback regarding their priorities for the new Advisory Committee on Data for Evidence Building by the deadline (February 9).
- NCHS seeks continued input from the BSC about the circumstances in which provisional data would be useful, despite the risk that early estimates may need to be revised later.
- Dr. Moyer welcomes suggestions from the BSC about how these meetings could be better structured.

- Future BSC meeting dates for the remainder of 2021: May 18-19; October 21-22.

Wednesday, January 27

Presenters

Brian Moyer, Ph.D., Director, NCHS
James Craver, M.A.A., Special Assistant to the Center Director, NCHS
Travis Hoppe, Ph.D., Chief Data Scientist, NCHS
Jennifer Parker, Ph.D., Director, Division of Research Methodology
Irma Arispe, Ph.D., Director, Division of Analysis and Epidemiology (DAE)
Renee M. Gindi, Ph.D., Chief, Population Health Reporting and Dissemination Branch, DAE
Florence Lee, M.P.H., Health Statistician, Population Health Reporting and Dissemination Branch, DAE
Stephen Blumberg, Ph.D., Director, Division of Health Interview Statistics
Ryne Paulose, Ph.D., Acting Director, Division of Health and Nutrition Examination Surveys (DHANES)
Cynthia L. Ogden, Ph.D., NHANES Analysis Branch Chief, DHANES
Tom Krenzke, Vice President and Associate Director, Westat
Carol DeFrances, Ph.D., Acting Director, Division of Health Care Statistics (DHCS)
John Lumpkin, M.D., M.P.H., Chair, NAMCS Workgroup

Welcome, Introductions, and Call to Order

Linette T. Scott, M.D., M.P.H., Chair, BSC
Sayeedha Uddin, M.D., M.P.H., Designated Federal Officer, NCHS, BSC

Dr. Uddin asked BSC members to introduce themselves and state any conflicts of interest. None of the BSC members stated a conflict of interest.

Dr. Scott called the meeting to order.

NCHS Update

Brian Moyer, Ph.D., Director, NCHS
James Craver, M.A.A., Special Assistant to the Center Director, NCHS

Dr. Moyer thanked the BSC for their many contributions. He noted that NCHS is near the end of the formal approval process for new BSC members. Although he cannot yet announce their names, he believes they will be approved before the next BSC meeting.

Dr. Madans retired at the end of December after years of exceptional service to NCHS. Dr. Moyer expressed his appreciation of her many valuable contributions over the years.

Recent Data on COVID-19

NCHS has aggressively shortened the timeline for releasing data related to COVID-19 and excess mortality. NCHS also introduced new data detailing the characteristics of mothers giving birth who tested positive for SARS-COV-2 and is devoting increased attention to racial equity.

NCHS Publications

Recently, NCHS released various publications (e.g., 2019 Mortality, Final Death Data for 2018).

New Advisory Committee on Data for Evidence Building

Dr. Moyer announced that a new committee was formed to assist the Office of Management and Budget in efforts related to evidence-based policy making—The Advisory Committee on Data for Evidence Building. Dr. Moyer serves on the committee along with others from academia, governmental agencies, and private organizations. The committee is currently developing a work plan and has made a request for public comment. Dr. Moyer encouraged the BSC to contribute feedback regarding their priorities; the deadline for submitting comments is February 9.

Budget Update

The enacted Fiscal Year (FY) 2021 budget (\$175M) represents a \$1M increase relative to the annual budget for FY2020. There was a realignment such that the \$14M NCHS usually received annually via the Centers for Disease Control and Prevention (CDC) is now directly and explicitly represented in the NCHS budget.

New CDC Director

CDC has a new director, Dr. Rochelle Walensky, who was formerly a Professor of Medicine at Harvard Medical School and a practicing physician at Massachusetts General Hospital. She is currently convening informational gatherings with CDC personnel. One of her top priorities is health equity.

CDC Data Modernization Initiative (DMI) Thematic Areas

Mr. Craver presented an overview of the DMI. NCHS is responsible for two of three thematic areas but is likely to contribute to the third area as well (i.e., new standards and approaches for public health reporting).

The first theme concerns data sharing across the public health ecosystem. Primary NCHS activities include modernizing the National Vital Statistics System (NVSS) and improving NVSS data quality and consistency.

The second theme pertains to enhancements related to ongoing data modernization. A key activity is the standardization of meta-data across all data systems, which will enable the development of new Application Programming Interfaces (APIs). Other activities in this domain comprise developing a new query system, renovating the NCHS website, and designing a virtual data enclave to allow remote access to restricted data.

Discussion/Reaction by the Board

One BSC member asked how the DMI relates to the activities of the committee on evidence-based policy. Dr. Moyer emphasized NCHS's commitment to providing policy-relevant data.

NCHS Strategic Planning

Brian Moyer, Ph.D., Director, NCHS
Travis Hoppe, Ph.D., Chief Data Scientist, NCHS
Jennifer Parker, Ph.D., Director, Division of Research Methodology

Dr. Moyer began by reiterating his key priorities: (1) harness alternative data sources; (2) extend the scope of analysis at NCHS; (3) connect data across the Center and statistical system; and (4) empower the NCHS workforce to implement this agenda. In terms of the operational

component, NCHS has adopted a two-pronged approach that includes case studies as well as formal strategic planning.

Dr. Moyer explained that further development of data science methods at NCHS is a work-in-progress. The update that Dr. Hoppe will provide is intended to solicit input from the BSC.

The State of Data Science at NCHS: Innovation Group

Dr. Hoppe, who joined NCHS as Chief Data Scientist about 3.5 months ago, began by addressing the question: What is data science? At NCHS, data science involves two components: process (i.e., scaling up, greater automation, ensuring sustainability) and people (i.e., collaboration across CDC, involvement of subject matter experts in all processes).

NCHS has three programs to facilitate the development of data science methods. The first of those (DMI) was already described by Mr. Craver. In the next segment, Dr. Parker will discuss the second program (Data Science Case Study). In this presentation, Dr. Hoppe focuses on the efforts of the innovation group to inspire and encourage data science projects. NCHS has held several staff-run innovation breakout sessions on various topics (e.g., artificial intelligence and machine learning; API implementation). Recently, the innovation group surveyed staff at NCHS who were engaged in work related to data science. More than 100 individuals representing 72 projects across all divisions within NCHS responded to the survey. The results demonstrate that data science is already active at NCHS and includes a wide variety of activities (e.g., enhancing data quality) and types of data output (e.g., restricted access files, public use files).

Dr. Hoppe highlighted some of the active projects. The redesign of *Health, US* involves transforming a book into an electronically accessible product that facilitates more timely reporting and provides small area estimation. The identification of opioid-involved health outcomes project combines multiple data sources including literal text. NVSS is using natural language processing to classify drowning deaths and to identify where the drowning occurred based on the literal text. NVSS is also using unstructured text to identify COVID-19 deaths that may not be directly identified as such. To provide more timely COVID-response data, NCHS is also using nowcasting and geospatial analysis.

Dr. Hoppe posed the following questions to the BSC: Given the risk that predictions may change, is it useful to provide provisional data? If so, in which cases?

All of these efforts are intended to enhance rather than replace traditional statistics. In closing, Dr. Hoppe highlighted new areas of growth: predictive analytics; use of machine learning for classification; remote access (i.e., API that will allow users to download and integrate data with other data systems); differential privacy (e.g., synthetic data); new datasets (e.g., retinal imaging); and bibliometric analysis to determine who uses/cites NCHS data.

Data Science Case Study Group

Dr. Parker explained that the mission of this group is to determine how well the work at NCHS aligns with what is needed. The Case Study Group promotes grassroots ideas by encouraging staff development and NCHS-wide engagement. The group also investigates ways to overcome the barriers. She asked the BSC: How should NCHS explain these new methods to the public and communicate that the results may be sensitive to assumptions and require later revision?

Discussion/Reaction by the Board

The discussion focused on the risk to NCHS's reputation, the importance of quantifying uncertainty, the need to balance competing priorities, overlap with similar work being done outside of NCHS, the value of synthetic datasets, and availability of training for staff.

A BSC member acknowledged the tradeoff between the benefits of disseminating preliminary data and the risk to NCHS's reputation. The board member suggested that NCHS might have a separate website dedicated to experimental data.

Many participants emphasized the need to incorporate measures of uncertainty. One BSC member noted that many machine learning algorithms do not incorporate uncertainty. Drs. Parker and Hoppe agreed that measures of uncertainty are crucial in the NCHS context.

BSC members asked how NCHS will balance development activities with traditional activities. Dr. Moyer acknowledged the inherent tradeoff between NCHS's traditional work and efforts to incorporate new technologies. The new methods involve risk, but if NCHS does not adapt to modern data needs, the data provided by NCHS is likely to become outdated.

A board member asked about potential overlap with government/private entities that may be engaging in similar development work. Drs. Hoppe and Parker emphasized that NCHS regularly consults with the wider HHS community and with colleagues in academia and elsewhere. Much of the previous work at NCHS has been in collaboration with participants outside NCHS.

One BSC member valued the fact that synthetic datasets could be made available to researchers who cannot access a Research Data Center (RDC). The member suggested consulting with key data users to determine what questions remain unanswered by synthetic data.

Another BSC member asked how NCHS is training staff to support the data science work? Dr. Parker noted that NCHS offers a lot of opportunities for training in many different forms (e.g., self-paced, groups). Drs. Parker and Hoppe both emphasized the value of staff development assignments (i.e., opportunities to learn by experience).

Health, United States, Redesign Update

Irma Arispe, Ph.D., Director, DAE

Renee M. Gindi, Ph.D., Chief, Population Health Reporting and Dissemination Branch, DAE

Florence Lee, M.P.H., Health Statistician, Population Health Reporting and Dissemination Branch, DAE

Dr. Arispe opened the session by explaining that several years ago, the BSC conducted a review and asked *Health, US* to incorporate input from external stakeholders to promote modernization.

Dr. Gindi explained that *Health, US* is a legislatively mandated, flagship publication that integrates data across NCHS and external private sources. The main goals of the redesign were to increase timeliness, its relevance to key stakeholders, and audience awareness/use. As part of the input phase (May 2019-May 2020), DAE is soliciting feedback by various methods (e.g., stakeholder interviews, web user survey, web analytics) to better understand what information users are seeking and which features users expect from a redesigned product.

A key question is: in what format should the data be distributed? PDF is useful for intense study, but not ideal for dissemination. HTML is better for immediate discovery and allows links to other relevant information. The current website reflects the expectation that print/PDF users progress through the content section-by-section. In contrast, a web-first product would be organized by topic. NCHS has already modified the *Health, US* website to more prominently feature a Data Finder tool that allows users to search for specific tables using filters. FastStats, which is HTML-based and thus more discoverable by search engines, also provides links to *Health, US* tables. Analysis of page views and downloads reveals that Data Finder views began to outpace other uses (e.g., PDF downloads, homepage views) by mid-2019 and spiked dramatically in March 2020 as attention focused on the pandemic.

Currently, NCHS is working on redesigning the *Health, US* website to be topic-based. It is important that the website integrates all datasets across and outside of NCHS.

How Do We Modernize Data Access?

Dr. Gindi explained that *Health, US* needs to be more flexible, dynamic, and interactive. To enhance interactivity, DAE is developing a tool for customized exploration (i.e., it will allow the user to select indicators of interest, select a subgroup of interest, view the data in table or chart format, and download the underlying data). This transformation requires translating the PDF trend tables into machine-readable datasets (e.g., comma-delimited files that can be easily imported into statistical software for further analysis). The data will be made available on open data platforms (e.g., data.cdc.gov).

Dr. Gindi closed the session by noting that while still in the input phase, DAE is also engaged in other dimensions of the redesign process (e.g., How can we produce content more efficiently? What content will be included? How can we disseminate data more effectively?).

Discussion/Reaction by the Board

The discussion centered on how *Health, US* differs from Healthy People, the variety of formats in which the data will be disseminated, and whether the website will allow open comments.

The BSC discussed the differences between Healthy People (which is prevention-oriented, with 10-year targets) and *Health, US* (which provides information regarding health that is not necessarily prevention-oriented such as health spending). Dr. Arispe noted that there is less overlap than one might think. However, NCHS is interested in exploring how these two products complement one another and could be better integrated.

In terms of the formats in which *Health, US* will be disseminated, a board member recognized the value of the machine-readable formats but hoped NCHS will continue to provide the data in PDF format. Dr. Gindi confirmed that the team is exploring the ability to offer both PDF and machine-readable files. The BSC member also asked whether NCHS will incorporate measures of uncertainty into the visualization tool. NCHS was already planning to do that for the machine-readable files and realizes that visualizing uncertainty is another area for the team to explore.

A participant asked whether the redesigned website will allow open comments, both positive and negative. Dr. Gindi confirmed that they have incorporated that feature into the mock-ups for the redesigned website.

Integrating Pulse and RANDS Surveys into the National Health Interview Survey (NHIS)

Stephen Blumberg, Ph.D., Director, Division of Health Interview Statistics
Jennifer Parker, Ph.D., Director, Division of Research Methodology

Dr. Blumberg began by expressing NCHS's desire to make survey data more timely but is seeking feedback from the BSC about these efforts.

Providing Near Real-Time COVID-19 Data with Experimental Online Surveys

Dr. Blumberg repeated a presentation he gave in November 2020 to CDC senior leaders because he wants to pose the same questions to the BSC. Using NHIS as an example, he admitted that it can take NCHS months or even years to modify the survey because they pre-test everything to ensure the survey remains robust. Given the time lag, the NHIS data are not very useful for surveillance. Although NHIS does have an Early Release Program, it includes only key indicators that are already on the survey and a minimum six-month delay remains between data collection and early release. When data that are more timely and responsive to changing needs are needed, an online survey may be a good option.

NCHS is already conducting two online surveys. In collaboration with Census, the Household Pulse Survey was launched in April 2020, has been in the field continuously since then, and the data are released every two weeks. Results show that reported levels of anxiety and depression increased during the spring and summer of 2020 and were substantially higher than in the first half of 2019 (based on NHIS). The survey also indicates that, among adults with less than a high school degree, the percentage who are uninsured increased in late-May of 2020.

Dr. Blumberg posed a hypothetical schedule in which data would be collected from both NHIS and an online survey during January through March, advance estimates based on the online survey would be released in May, the NHIS early release data would be released in September as usual, and the final estimates based on NHIS for the entire year would be released in June of the following year. He asked the BSC whether that schedule would be useful for some topics (e.g., mental health, insurance coverage)? He warned that the advance estimates are not likely to be very reliable given the low response rate (~7%) to the Pulse survey. The advance estimates are unlikely to match the final estimates.

The second online survey is the Research and Development Survey (RANDS), which uses probability-based online survey panels and includes incentives (i.e., money or prizes) to enhance participation. Among people who agreed to be in the panel, the Round 1 completion rate was 78%. However, among all people asked to participate, the response rate was 23%. Comparisons between the estimates from Round 1 of RANDS and NHIS reveal that in some cases they are similar (e.g., uninsured, diabetes prevalence), but in other cases they differ by a lot (e.g., couldn't afford care, delayed care). Online surveys could be particularly useful for addressing emerging topics. For example, RANDS released data on use of telemedicine in August 2020, about one year before NHIS will be able to release such data.

NCHS is exploring how supplemental online surveys might enhance the traditional surveys. Online surveys could be used for advance question development; to address emerging topics; to provide more targeted, detailed questions; and to facilitate more timely estimates. In November 2020, Dr. Blumberg posed two questions to CDC senior leaders: (1) How can NCHS effectively communicate the differences in data quality between different survey types? (2) Are online surveys the right strategy to produce more timely and responsive data? CDC senior leaders agreed that NCHS needs a mechanism for more timely data, but expressed concern

about the data quality from online surveys. Such surveys may be useful for some topics but not others.

Proposed: Commercial Panels for Online Quarterly Data Collection

During the two months since that meeting, Dr. Parker explained that NCHS has developed a proposal to use online surveys to publish quarterly estimates with only a one to two month delay. However, more methodological research is needed to resolve several issues (e.g., comparisons of direct estimates with NHIS, methods for calibrating estimates to established benchmarks, how to use RANDS for nowcasting/forecasting models, head-to-head panel comparisons).

Discussion/Reaction by the Board

The discussion was organized around questions related to four themes outlined by Dr. Parker.

Quality of online survey estimates

There was general consensus among the board that online surveys are likely to provide better information regarding trends than levels. One BSC member suggested that traditional and online surveys can be treated as parallel streams of data that serve different purposes and need not be synchronized. Another board member noted that quarterly release may not be frequent enough in the midst of a crisis. This person also advocated for including confidence intervals to encourage users to view the estimates as a range. Comparing estimates from multiple panels could provide a measure of robustness.

How to prioritize efforts to improve timeliness

One board member thought that the NHIS early release estimates should have higher priority than advance estimates from online surveys. Despite the delay, most other government surveys do not offer anything as quickly as the NHIS early release. Another BSC member asked why NCHS must prioritize one over the other? The problem, as Dr. Blumberg explained, is limited resources in terms of money and staff. A third board member thought it would be catastrophic if NCHS released an advance estimate based on the online survey, but the NHIS estimate was totally different.

Identification of key health outcomes (for what topics does it make sense?)

Which topics are more important for early estimation may depend on the current issues. One BSC member suggested that even if the estimates are wrong, they are likely to be stable because they are based on samples of 100,000 respondents every two weeks. Thus, why not use the NHIS to obtain the point estimates, but then use data from the online survey to extrapolate the more recent trend? Another board member endorsed that idea, noting that it would avoid big discrepancies between the estimates based on online surveys versus NHIS.

Communicating quality

A board member suggested that NCHS could follow one of the models used by other federal statistical agencies. For example, some agencies publish completely separate series for quarterly and annual estimates. The Bureau of Labor Statistics uses a different model in which it publish initial estimates, which are later revised.

NHANES Update

Ryne Paulose, Ph.D., Acting Director, DHANES
Cynthia L. Ogden, Ph.D., NHANES Analysis Branch Chief, DHANES
Tom Krenzke, Vice President and Associate Director, Westat

Dr. Paulose noted that NHANES is involved in many other activities (e.g., redesigning NHANES for 2023), but today's presentation will focus on the 2019-20 Pre-Pandemic data.

Pre-Pandemic 2017-20 Sample Weights

Dr. Ogden explained that NHANES was halted because of the pandemic, and thus, only 18 of the 30 PSUs were completed in the 2019-20 survey cycle. In addition, 2019 was the first time that data from a single year were not nationally representative. Consequently, NCHS decided that the partial data for 2019-20 would need to be combined with 2017-18. Developing sample weights for this 2017-20 Pre-Pandemic dataset requires combining data from two different sample designs. The 18 completed PSUs from 2019 and early 2020 were mapped to the 2015-2018 sample design strata. A PSU adjustment factor of 4/divided by the number of PSUs in each stratum was applied to all sample weights. For example, if the stratum included 2 PSUs, the adjustment factor was 2.0—thereby upweighting the PSU, whereas if the stratum included 6 PSUs, the adjustment factor was 0.67—thus, down-weighting the PSUs.

To evaluate the sample weights, DHANES compared the weighted estimates with estimates from the American Community Survey (ACS), NHIS, and previous NHANES cycles. The first set of weights (W1) revealed differences between NHANES and ACS in the distributions by population size and urban/rural. A second set of weights (W2) was created by raking to four population size groups. The W2 weights matched ACS well in terms of population size, but sizable discrepancies in the urban/rural distribution remained. For the third set of weights (W3), DHANES raked to four rural-urban categories instead of population size. The W3 weights matched ACS better in terms of urban/rural, but the distribution by population size differed.

DHANES then compared estimates for selected variables (e.g., education) between NHANES and ACS. Although there were some differences, the results were similar using W2 versus W3, and they were comparable with the disparities between previous NHANES cycles and ACS. Next, DHANES compared estimates for other variables (e.g., self-reported health status) with NHIS. Again, the discrepancies were consistent over time and similar using W2 versus W3. Finally, comparisons of 2017-20 NHANES with previous cycles of NHANES suggested that the estimated percentages of adults with high total cholesterol and obesity appeared consistent with previous trends, regardless of whether weights W2 or W3 were used.

As a next step, DHANES plans to simulate the weight adjustment using 2013-16 incomplete datasets. DHANES is leaning toward using W3 rather than W2 because health disparities by urban vs. rural are generally greater than disparities by population size.

Dr. Ogden emphasized that 2019-20 should not be separated from the 2017-18 data because the 2019-20 data are not nationally representative. Therefore, DHANES will make some changes to the files (e.g., renumbering sequence identification numbers) to discourage identification of which records came from 2019-20 versus 2017-18.

She closed by posing two main questions to the BSC. First, how should NCHS communicate that stand-alone analysis of 2019-20 is not possible? Second, how should NCHS represent the time period given that the data collection was pre-pandemic and excludes most of 2020?

Disclosure Risk Assessment

Mr. Krenzke explained that because only 18 PSUs were completed in 2019 to March 2020 pre-pandemic versus 30 PSUs in 2017-18, there is increased risk that a particular respondent from 2019 to March 2020 pre-pandemic could be identified. DHANES has already made some data modifications to reduce that risk (e.g., recoding selected variables, suppressing other variables). Using this revised dataset, the Westat team evaluated the disclosure risk based on several assumptions (e.g., the intruder knows which 18 counties were sampled in 2019 to March 2020 pre-pandemic and can identify the records from 2017-18 versus 2019 to March 2020 pre-pandemic by differencing the 2017-20 public-use file against the 2017-18 public-use file). The results indicate that 8 counties can be easily identified, from which another 6 counties can logically be re-identified. The remaining 4 counties have a lower chance of identification. The team modeled six alternative scenarios assuming different sets of variables are identified. The risk of identifying a particular respondent was low-to-moderate in two of the scenarios: (1) the intruder identifies 14 counties and (2) the intruder identifies 8 counties. It is also assumed that the intruder has information about nine other identifying variables (i.e., sex, age, race/ethnicity, country of birth, education, marital status, ever served in armed forces, number of children aged 5 or younger, household income). The disclosure risk is reduced if the intruder does not know household income, whether the person ever served in the armed forces, or the number of young children in the household.

Based on these results, Mr. Krenzke recommends several confidentiality edits. First, some information (e.g., served in the military, household income) should be suppressed and made available only via the RDC. Other variables (i.e., marital status, length of time in the United States) should be recoded to combine categories with low frequency. After making those changes, he recommends that Westat re-run the risk assessment analysis to determine whether it is necessary to further suppress the respondent's age in years and pregnancy status. The usual practice is to mask the variance estimation codes via controlled random swapping. Mr. Krenzke recommends increasing the swapping rate for PSUs that are easily identified and for individuals with high-risk of identification based on the model.

Timeline for Data Release

Dr. Paulose explained that DHANES plans to release a 2017-20 Pre-Pandemic dataset that combines the 2017-18 and 2019-20 cycles to form a four-year nationally representative sample. The dataset will include only those components collected in both cycles. Components that were introduced in 2019-20 will only be available through RDC and treated as a convenience sample.

DHANES expects the disclosure review board will approve the plan in February, after which it must undergo collaborator review. Dr. Paulose expects NCHS will be able to release the public-use file in April. As in the past, the first wave of data release will focus on high-priority information (i.e., demographic variables, sample weights, selected variables from the questionnaire, exam, and laboratory assays). The data release will also include an analytic document that outlines appropriate use of the data. The full dataset will be available from the RDC.

Discussion/Reaction by the Board

Themes highlighted in the discussion included how to frame the dataset, the possibility of adjusting the weights for both population size and urban/rural, and whether the risk of disclosure may be over-estimated because it assumes no individual variability in response.

Multiple board members endorsed the label “2017-20 Pre-Pandemic,” which emphasizes that the data represent the “pre-pandemic” period. One person noted that surveys often include the years in the title even when the fieldwork does not cover an entire year.

One BSC member asked whether DHANES had considered adjusting the weights for both population density and urban/rural status. This person also wondered whether DHANES has evaluated whether there was any change between 2017-18 and 2019-20 within strata. Dr. Ogden replied that they have not, but will do so.

Another board member asked Mr. Krenzke whether the analysis accounted for individual variability in responses over time. Mr. Krenzke explained that they assume the information provided by the respondent represents the “truth.” The BSC member replied that, in that case, they may be over-estimating the disclosure risk. Mr. Krenzke agreed that their estimates are conservative.

National Ambulatory Medical Care Survey Update

Carol DeFrances, Ph.D., Acting Director, DHCS
John Lumpkin, M.D., M.P.H., Chair, NAMCS Workgroup

NAMCS BSC Workgroup Report

Dr. DeFrances began by thanking the BSC workgroup for their excellent work. The BSC workgroup report provided DHCS with an overarching framework that helps inform the NAMCS redesign. DHCS has since created a NAMCS redesign workgroup to address the major questions raised in the report.

2020 NAMCS and COVID-19

NAMCS made some changes in 2020 to reduce the burden on providers and to maximize safety. Fielding of the physician component was delayed by 6 weeks (i.e., to the end of May) and by an additional 4.5 weeks in COVID-19 hot spots. Interviews were conducted primarily by phone rather than in-person. NAMCS did abstraction of visits during the first quarter, but the response was very low. Consequently, NAMCS conducted only the induction interviews after the first quarter. The overall response rate was 40-50%.

The Community Health Center (CHC) component, which usually starts during the summer, began on time. Interviews were conducted by phone. Unlike the physician component, NAMCS was able to continue abstraction over the entire year. Although the response rate was lower than in prior years, Dr. DeFrances believes it will be sufficient to support national estimates.

COVID-19 related questions (i.e., shortages of personal protective equipment, ability to test/refer to a testing site, need to turn away virus-positive patients, providers who tested positive, use of telemedicine) were added in both physician and CHC components. Preliminary estimates for the physician component will be released as soon as possible.

2021 NAMCS

For 2021, the physician component will continue to be limited to physician interviews with no collection of visit data. In the latter half of the fieldwork period, NAMCS will expand the universe of physicians targeted to include some types of providers that were previously ineligible. DHCS is also collaborating with the U.S. Census Bureau to explore alternative sampling frames (e.g., business-level databases) and to identify supplemental sources for obtaining visit-level data.

A big change to the CHC component is that NAMCS will begin to collect electronic health records (EHRs). To help defray the cost of implementation, NAMCS will offer a \$10,000 set-up fee to the CHC. Dr. DeFrances announced that DHCS was awarded a Patient-Centered Outcomes Research Trust Fund grant for FY21 to evaluate the linkage of maternal health EHR data with the National Death Index and the Housing and Urban Development administrative data.

2022 NAMCS

DHCS is already planning ahead for 2022. The physician interview will be revised to capture more relevant ambulatory care topics and to reduce interview length to enhance response rate. DHCS is also exploring whether to separate the induction interview from the collection of data on visits. The CHC component will continue collecting EHR data but will work to expand the sample to 110 Federally Qualified Health Centers or similar health centers. DHCS sent a list of questions to the BSC Workgroup, the responses to which Dr. Lumpkin will discuss next.

NAMCS Workgroup

Dr. Lumpkin began by acknowledging the difficult task faced by NAMCS because of transitions in ambulatory care (e.g., many different kinds of providers, care being delivered in various sites, multiple payment schemes, substantial increase in telemedicine). He then summarized the NAMCS Workgroup's opinion with its responses to the questions posed by the NAMCS team.

Should other types of advanced practice providers be included?

The BSC Workgroup suggested including physician assistants, nurse practitioners, certified nurse-midwives, and all types of advanced practice nurses. However, it recognizes that it will be a challenge to identify all the various providers.

How do we include them in a way that is nationally representative?

Given that there is no inclusive national database for nurses providing direct primary care, NAMCS will need to link together many datasets to try to get enough of a database to design a sampling frame that will cut across the different ways that ambulatory care is delivered by nurses in the United States.

What databases for advanced practice providers are available?

Dr. Lumpkin provided a long list of databases from which a sampling frame could be drawn but acknowledged the potential for substantial redundancy across these databases.

Should NAMCS incorporate alternative databases for physicians?

The BSC Workgroup suggested continued use of the American Medical Association and American Osteopathic Association master files for NAMCS sampling because they are considered the gold standard. However, information from individual certifying boards could be used for confirmatory purposes.

If NAMCS needs to use multiple databases to draw sample, how should they be combined?

The BSC Workgroup does not have the expertise to answer that question, but suggested that DHCS consult with database super users (e.g., Bob Berenson at the Urban Institute, Stephen Petterson at the Robert Graham Center, Eugene Rich at Mathematica) to determine the best methods for combining those files.

Thoughts regarding separating physician interview from visit data?

The BSC Workgroup believes that NCHS will learn a lot from its experience collecting EHR data in CHCs. It also endorsed the use of more phone-based and remote interview methods (e.g., Zoom).

Discussion/Reaction by the Board

The BSC expressed general support for the plans outlined for NAMCS but emphasized the need for appropriate due diligence. One BSC member wondered whether there were consent issues with the use of EHR data. Dr. DeFrances explained that previously NAMCS had obtained a waiver of informed consent for patients and physicians. However, due to changes to the Common Rule, NAMCS is now considered public health surveillance rather than research. She was not sure of the impact of this change on NAMCS.

Actions

Dr. Scott called for a vote to accept the final report from the BSC NAMCS workgroup and for the BSC to formally issue recommendations based on the suggestions in the workgroup's opinion report. The vote was unanimous in support among attending BSC members (Dr. Olson was no longer present).

BSC Wrap-up

Linette T. Scott, M.D., M.P.H., Chair, BSC
Brian Moyer, Ph.D., Director, NCHS

Dr. Moyer thanked everyone for their participation. Strategic planning was a common theme throughout this meeting. The presentations were intended to stimulate discussion. He expects today to be the first in a series of discussions. NCHS would like to engage the BSC throughout the process so that the BSC can help inform the agenda. Dr. Moyer also welcomes suggestions from BSC members about how these meetings could be better structured.

At the last meeting, the BSC agreed to draft a formal letter to HHS, CDC, and NCHS leadership. Dr. Scott apologized and indicated that the letter is still in process. She plans to finalize the letter and send it to the BSC for review soon.

Public Comment

There was no public comment. Dr. Uddin explained that public comments can also be submitted via post or can be directed to Dr. Uddin via email.

The meeting was adjourned at 4:50 pm ET.

To the best of my knowledge, the foregoing summary of minutes is accurate and complete.

_____/s/_____
Linette T. Scott, M.D., M.P.H.
Chair, BSC

4/26/2021
DATE