

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
May 26, 2022**

Meeting Summary

The Board of Scientific Counselors (BSC) convened via Zoom on May 26, 2022. The virtual meeting was open to the public (via Zoom).

Board Members Present

John R. Lumpkin, M.D., M.P.H., Chair, BSC
Mollyann Brodie, Ph.D.
Kennon R. Copeland, Ph.D.
Tara Das, Ph.D., M.P.H., M.L.I.S.
Scott H. Holan, Ph.D.
V. Joseph Hotz, Ph.D.
Helen G. Levy, Ph.D.
Bradley A. Malin, Ph.D.
Sally C. Morton, Ph.D.
Lucila Ohno-Machado, M.D., Ph.D.
Kristen M. Olson, Ph.D.
Andy Peytchev, Ph.D.
C. Matthew Snipp, Ph.D.
Kelly Hoover Thompson, J.D.
David R. Williams, Ph.D., M.P.H.

CDC/NCHS Panelists

Brian Moyer, Director, National Center for Health Statistics (NCHS)	Askari Rizvi Stephen Blumberg James Craver	Gwen Mustaf Kristen Olson Andy Peytchev
Rebecca Hines, Designated Federal Official, NCHS, BSC	Carol DeFrances Renee M. Gindi Dan Jernigan	J. Neil Russell Paul Sutton Brian Ward
Irma Arispe	Lisa Mirel	

Public Attendees

Joyce Abma	Johanna Alfier	Francis Barlas
Nikki Adams	Robert Anderson	Bryan Bassig
Dzifa Adjaye-Gbewonyo	Hallie Andrews	Lindsey Black
Viviana Aguila	Nicholas Ansai	Jonaki Bose
Farida Ahmad	Brenda Baker	Richard Boyd
Lara Akinbami	Kai Baker Stringfield	Kate Brett
Juan Albertorio	Tammy Banks	Beckie Brooks

Amy Brown
Cynthia Bush
Eileen Call
David Carranza
Nora Castro
Amy Cha
Anjani Chandra
Te-Ching Chen
Leah Christian
Robin Cohen
Traci Cook
Christine Cox
Lauren Creamer
Nicole Cummings
La-Tonya Curl
Doug Currivan
Doug Currivan
James Dahlhamer
Barnali Das
Orlando Davy
J. Michael Dennis
Avay Dolberry
Annette Durr
Morgan Earp
Nazik Elgaddal
James Ferguson
Steven Fink
Sheila Franco
Alicia Frasier
Cheryl Fryar
Adena Galinsky
Matthew Garnett
Cordell Golden
Vanessa Green
Lello Guluma
Leda Gurley
Heidi Guyer
Brady Hamilton
Nancy Han
Althelia Harris
Yulei He
Elizabeth Heitz
Melonie Heron
Kevin Heslin
Jacquie Hogan
Isabelle Horon
John Hough
Rebecca Hu

David Huang
Marisol Iniguez
Katherine Irimata
Geoff Jackson
Elizabeth Jackson
Lenel James
William Johnson
Jessly Joy
David Justice
Ashley Kaiser
Jessica Keralis
Gina Kilpatrick
Jieun Kim
Richard Klein
Ellen Kramarow
Nataliya Kravets
Denys Lau
Florence Lee
Jeffrey Leintz
Jessica Lendon
Sarah Lessem
Xianfen Li
Anthony Lipphardt
Michelle Liu
Denise Love
Jacqueline Lucas
Crescent Martin
Joyce Martin
Michael Martinez
Meredith Massey
Brian McGough
Jody McLean
Grace Medley
Justin Mezetin
Diane Michel
Leyla Mohadjer
Mary Moien
jennifer Moore
Kelly Myrick
Zakia Nelson
Amanda Ng
Tina Norris
Susan Nowlin
Colleen Nugent
Tatiana Nwankwo
Dave Nyczepir
Titilayo Okeyode
William Olesiuk

Sarah Osborne
Michelle Osterman
Cedric Palmore
Shaundell Pannell
Jessie Parker
Priyam Patel
Sue Pedrazzani
Whitney Perkins Witt
Zachary Peters
Michael Peters
Steve Pierson
John Pleis
Anthony Quintana
Catherine Rappole
Daniela Relf
Dean Resnick
Cynthia Reuben
Dorothy Roper
Cheryl Rose
Lauren Rossen
Charles Rothwell
Valerie Ryan
Brian Salant
Loredana Santo
Neda Sarafrazi
Jennifer Sayers
Paul Scanlon
Susan Schechter
Jeannine Schiller
Bobbie Shimizu
Merianne Spencer
Marietta Squire
Renee Storandt
Jane Sudol
Makram Talih
Liviu Tanase
Alexander Tin
Jianmin Xu
Jiaquan Xu
Jing Xu
Alana Yick
Henry Yin
Benjamin Zablotsky
Yasir Zafar
Cindy Zhang
Julie Zimmer

Meeting Tech
Meeting Writer

Mike Kavounis (RLA)
Christina Huffman (RLA)

List of Abbreviations

API	application programming interface
ASPE	HHS Office of the Assistant Secretary for Planning and Evaluation
ATO	authority to operate
BSC	Board of Scientific Counselors
CCQDER	Collaborative Center for Questionnaire Design and Evaluation Research
CDA	Consolidated Document Architecture
CDC	Centers for Disease Control and Prevention
CIPSEA	Confidential Information Protection and Statistical Efficiency Act
COMEC	Coordinating Office for Medical Examiners and Coroners
CMS	Centers for Medicare and Medicaid Services
CSP	cloud service provider
DAE	Division of Analysis and Epidemiology
DHCS	Division of Health Care Statistics
DMI	Data Modernization Initiative
DOS	Department of State
DQS	Data Query System
DUA	data use agreement
eCR	electronic case reporting
EDAV	Enterprise Data Analytics and Visualization
EHR	electronic health record
FHIR	Fast Health Interoperability Resources
FQHC	federally-qualified health center
FQHC-LAL	FQHC look-alike (HRSA designation)
FSRDC	Federal Statistical Research Data Centers
FY	fiscal year
GRASP	Geospatial Research, Analysis, and Services Program
HEHR	Healthcare Electronic Health Record (CDC system)
HHS	Department of Health and Human Services
HL7	Health Level Seven
HRSA	Health Resources and Services Administration
ICD-10	International Classification of Diseases version 10 (international coding standard)
IG	implementation guide
MDI	Medicolegal Death Investigation
ML	machine learning
NAMCS	National Ambulatory Medical Care Survey
NAPHSIS	National Association for Public Health Statistics and Information Systems
NCHS	National Center for Health Statistics
NCVHS	National Committee for Vital and Health Statistics
NDI	National Death Index
NHAMCS	National Hospital Ambulatory Medical Care Survey
NHANES	National Health and Nutrition Examination Survey

NHCS	National Hospital Care Survey
NHIS	National Health Interview Survey
NIOSH	National Institute of Occupational Safety and Health
NLP	natural language processing
NPALS	National Post-Acute and Long-term Care Study
NSDS	National Secure Data Service
NSFG	National Survey of Family Growth
NVSS	National Vital Statistics System
OMB	Office of Management and Budget
ONC	Office of National Coordinator
PHSPMDP	Population Health Survey Planning, Methodology and Data Presentation Workgroup
PHWG	Public Health Working Group
PII	personally-identifiable information
PPRL	Privacy Preserving Record Linkage
RDC	Research Data Center
SDOH	social determinants of health
STEVE	State and Territorial Exchange of Vital Events
STLT	state, tribal, local, and territorial
TEFCA	Trusted Exchange Framework and Common Agreement
T-MSIS	Transformed Medicaid Statistical Information System
USCDI	United States Core Data for Interoperability
VA	Department of Veteran Affairs
VDE	Virtual Data Enclave
VPN	virtual private network
VRDR	Vital Records Death Reporting
WONDER	Wide-ranging Online Data for Epidemiologic Research system

Action Steps

- The BSC voted unanimously to approve the report of the Population Health Survey Planning, Methodology and Data Presentation (PHSPMDP) Workgroup regarding web-based panel surveys.

Presenters

John Lumpkin, M.D., M.P.H., Chair, BSC, NCHS

Rebecca Hines, M.H.S., Designated Federal Officer, BSC, NCHS

Brian Moyer, Ph.D., Director, NCHS

Dan Jernigan, M.D., M.P.H., Deputy Director for Public Health Science and Surveillance, NCHS

James Craver, M.A.A., Deputy Director for Management and Operations, NCHS

Paul Sutton, Ph.D., Deputy Director, Division of Vital Statistics, NCHS

Carol DeFrances, Ph.D., Acting Director, Division of Healthcare Statistics, NCHS

Brian Ward, Ph.D., Chief, Ambulatory and Hospital Care Statistics Branch, Division of Healthcare Statistics, NCHS

Askari Rizvi, M.S., M.B.A., P.M.P., Chief, Technical Services Branch, Division of Healthcare Statistics, NCHS

J. Neil Russell, Ph.D., Director, Research Data Center, NCHS

Renee M. Gindi, Ph.D., Chief, Population Health Reporting and Dissemination Branch, Division of Analysis and Epidemiology, NCHS

Lisa Mirel, M.S., Chief, Data Linkage Methodology and Analysis Branch, Division of Analysis and Epidemiology, NCHS

Irma Arispe, Ph.D., Director, Division of Analysis and Epidemiology, NCHS

Andy Peytchev, Ph.D., BSC Member and Chair, Population Health Survey Planning, Methodology and Data Presentation (PHSPMDP) Workgroup

Kristen Olson, Ph.D., BSC Member and Chair, Workgroup to Consider and Assess Measures of Discrimination for Use in NCHS Surveys

Welcome and Call to Order

John Lumpkin, M.D., M.P.H., Chair, BSC, NCHS

Rebecca Hines, M.H.S., Designated Federal Officer, BSC, NCHS

Dr. Lumpkin opened the meeting by welcoming BSC members and NCHS staff.

Ms. Hines conducted roll call, asking members to introduce themselves and state conflicts of interest. None of the BSC members stated a conflict of interest. Dr. Peytchev noted that he is a fellow at RTI, which supports the National Survey on Family Growth (NSFG). Ms. Hines also recognized the efforts of the members of the two BSC workgroups that were presenting at the meeting.

NCHS Director's Update

Brian Moyer, Ph.D., Director, NCHS

Board, Staffing, and NCHS Office Updates

Dr. Moyer noted that the theme of this meeting is the NCHS Data Modernization Initiative (DMI), which is part of the larger Centers for Disease Control and Prevention (CDC) DMI. As part of this theme, the meeting included a presentation by Dr. Jernigan on DMI activities across CDC and how NCHS's efforts support these broader activities, which play a key role in modernizing public health infrastructure and support many of CDC's strategic initiatives.

Dr. Moyer thanked Drs. Olson and Levy for agreeing to extend their time on the BSC until the nomination process for new BSC members is completed.

Dr. Moyer highlighted that the Office of Management and Budget (OMB) recently named Dr. Karin Orvis, Ph.D. as the Chief Statistician of the United States. This post was previously vacant for 2 years, during which time many member agencies of the Federal Statistical System have provided interim leadership and support. Dr. Orvis will oversee the Federal Statistical System, including implementation of the Evidence-Based Policymaking Act. Dr. Moyer proposed inviting Dr. Orvis to a future BSC meeting to discuss upcoming changes to the Federal Statistical System.

Dr. Moyer described how NCHS has started to welcome back staff to its physical office, which has been renovated to improve ventilation and add other amenities. NCHS will continue to offer workplace flexibility as it transitions to a hybrid work environment.

Evidence-Based Policymaking Act Implementation

Dr. Moyer provided updates on implementation of the Evidence-Based Policymaking Act, particularly the requirement for a standard application process for researchers seeking to access restricted data across multiple federal agencies. Currently, researchers must submit separate applications to each agency to gain access to restricted data. The Act seeks to streamline this approach with one standard application across federal agencies while maintaining proper privacy restrictions within each agency.

OMB recently published a *Federal Register* notice that outlined a potential approach to this standardized application process. To support this process, NCHS is building a catalog of restricted datasets available at CDC research centers, which will be available on the CDC website to enable researchers to quickly identify relevant datasets. NCHS is also updating and correcting CDC metadata to ensure that datasets have proper labels, descriptions, and citations. These updates will be vital to implementing the standard application process outlined by OMB.

The Advisory Committee on Data for Evidence Building advises OMB on implementing the Evidence-Based Policymaking Act. The Advisory Committee is currently exploring development of a National Secure Data Service (NSDS), which would facilitate data sharing and linkages across federal, state, and local entities for evidence-based policymaking. The Advisory Committee is developing multiple use cases for the NSDS, including a use case for modernizing the National Vital Statistics System (NVSS). NCHS is currently supporting these use cases to not only modernize NVSS but also illustrate NCHS's role in establishing and maintaining strong data standards and data governance frameworks. Modernizing the NVSS as part of the NSDS will promote greater interoperability between federal, state, and local entities.

NCHS Budget Update

In fiscal year (FY) 2022, Congress increased NCHS's budget by \$5.5 million to support DMI activities, including development of next generation surveys and other NCHS modernization initiatives. The Biden administration's FY2023 budget request includes an increase of \$6.54 million for NCHS to facilitate larger sample sizes for the National Health Interview Survey (NHIS) and evaluations of current race and ethnicity categories in NCHS surveys. The final FY2023 budget will be determined by Congress in the coming months.

NCHS Program Updates

Dr. Moyer provided updates on multiple NCHS programs:

- The CDC Collaborative Center for Questionnaire Design and Evaluation Research (CCQDER) supported the U.S. Department of State's (DOS) efforts to add an option of an "X" gender marker

on passport applications for transgender and nonbinary people. CCQDER provided cognitive testing support that was vital to this effort.

- The NCHS team that supports international civil registration and vital statistics helped launch the 2022 World Health Organization Verbal Autopsy Instrument, which compiled more than 28,000 verbal autopsy records from 13 countries. Compiling verbal autopsy records is crucial for identifying gaps in mortality measurements such as deaths that occur outside of health care settings.
- The Division of Health Care Statistics (DHCS) is collaborating with the U.S. Census Bureau on the Retail Health Clinic Project. This project gathers data on retail health clinics and their surrounding geographic areas as a measure of health care accessibility and markets in different areas.
- NCHS will begin granting Virtual Data Enclave (VDE) access to researchers on select CDC-based projects. The VDE will streamline the process for researchers to access restricted CDC data and promote greater data-sharing among CDC offices and programs.

In addition to these programs, CDC recently established the Coordinating Office for Medical Examiners and Coroners (COMEC) office, which will serve as an information hub for medical examiners and other federal, state, and local stakeholders. Medical examiners and coroners provide vital public health data (e.g., cause of death), and the COMEC office will strengthen CDC's cooperation with medical examiners and coroners across the country. COMEC will also work to improve data collection, automation, distribution, and validation of death investigation data.

NCHS recently released multiple reports and data such as 2021 Provisional Birth Data; Sexual Orientation Differences in Access to Care and Health Status, Behaviors, and Beliefs; and Maternal Mortality Rates in the United States, 2020.

NCHS collaborates with AcademyHealth to manage its health policy fellowship program, including the 2021 data visualization competition for undergraduate students. The competition focused on developing interactive data visualization using NCHS datasets to illustrate impacts of policies on important health trends or disparities. The first-place winner was "Sound the Alarm: Data Show Large Inequities in Rates of Gun Deaths Among Young People" by students from Johns Hopkins University, and second place was awarded to "Investigating Temporal Trends in Preventable Death: Suicide and Drug Overdose" by students from University of Wisconsin-Madison. The 2022 competition will allow contestants to use additional NCHS datasets while continuing to focus on how social determinants of health (SDOH) influence health outcomes.

NCHS will host its 2022 NCHS Data Detectives Camp in August, which is a summer camp for sixth- and seventh-grade students to learn about statistical concepts through interactive activities. The COVID-19 pandemic led NCHS to switch this camp from in-person to virtual, which enabled a larger number of campers across the United States to participate. Applications will be accepted through May 30.

Data Modernization Initiative (DMI) Updates and Discussion

CDC Update

Daniel Jernigan, M.D., M.P.H., Deputy Director for Public Health Science and Surveillance, CDC

The ultimate goal of CDC's DMI efforts is to move public health data infrastructure from siloed and rigid public health data systems to a connected, adaptable, flexible, and response-ready architecture that provides faster and more accurate data access for public health decision-making. CDC's DMI efforts aim to address multiple challenges posed by current public health data systems such as the following:

- Siloed information, including disconnected or proprietary systems driven by disease-specific budgets
- Need for greater training on current technologies by public health workforce
- Heavy reporting burden on providers, including sending health data to multiple agencies in different formats
- Older technologies at many health departments that are not flexible, cloud-based, or scalable
- Separation of public health departments and researchers from health care data ecosystem (e.g., electronic health records [EHRs]) and lack of federal incentives and regulations for public health data access

CDC worked with many different partners to establish five priorities that will guide DMI efforts, budgeting, and priority teams:

- **Build the right foundation** for a secure and accessible cloud-based system for key health data such as vital statistics, immunization data, lab reports, and case reporting from EHRs.
- **Accelerate data into action** for rapid responses to outbreaks and other public health emergencies by integrating data from numerous sources, including nontraditional data sources.
- **Develop a state-of-the-art workforce** at CDC that includes data scientists. This priority will be supported by a \$3 billion grant assigned through the American Rescue Plan for a five-year workforce improvement program.
- **Support and extend partnerships** with CDC partners, including numerous private industry partners, academia, and state, tribal, local, and territorial (STLT) partners.
- **Manage and change data governance**, including data integration and harmonization processes.

Implementing these DMI priorities requires changing CDC's approach to data sharing, new evidence requirements, and data management policies. CDC data are currently siloed in different programs such as Infectious Diseases and Science & Surveillance. This structure will change under the DMI Implementation approach, which charges five cross-agency teams to address each of the five DMI priorities listed above. These cross-agency teams will ensure consistent approaches to DMI and data governance across agencies as well as greater data sharing and coordination.

To support data modernization, CDC established Enterprise Data Analytics and Visualization (EDAV), a centralized cloud-based data analytics and visualization platform that can be used across CDC agencies. Centralizing data access and analytics will break down data siloes and streamline data analyses and research by CDC scientists. EDAV is built in Microsoft Azure, and CDC is currently connecting EDAV with other health statistic systems such as the STLT Data Hub and Cerner Health Data Lab. EDAV has already enabled CDC to automate reporting for the Pertussis program. This automation resulted in a 95 percent reduction in reporting time compared to the previous manual reporting process.

CDC is collaborating with the Office of National Coordinator (ONC) to plan data modernization approaches. CDC and ONC representatives meet biweekly to discuss and plan approaches for meeting 21st Century Cures Act deadlines, Centers for Medicare & Medicaid Services (CMS) inpatient prospective payment system regulations, and other DMI requirements. Over the next 2 years, CDC and ONC will collaborate to establish the North Star Architecture as a centralized public health data infrastructure for data sharing between CDC and STLT partners. CDC is also changing how it engages with partners to further facilitate data sharing, including industry listening sessions, a joint Data Surveillance Workgroup, and a Consortium for Data Modernization.

Modernizing public health data infrastructure and tools can provide numerous benefits, including (1) decreased reporting burdens on hospitals and doctors; (2) freeing of public health staff from manual analyses so that they can focus more on targeted interventions; and (3) greater awareness of emerging

public health threats, improving allocation of resources to mitigate public health impacts. The COVID-19 pandemic illustrated the sheer amount of data available for public health analyses and decision-making; these data include 813 million electronic lab reports for SARS-CoV-2 tests, 551 million vaccination records, and 140 terabytes of clinical and administrative data. Integrating such large datasets across public health agencies will significantly improve public health analyses and approaches.

Discussion/Reaction by the Board

CDC selected Microsoft Azure as the cloud service provider (CSP) for EDAV based on a cost and efficiency determination, but CDC plans for a data architecture that can work across other CSPs (e.g., Amazon Web Services [AWS]) and data platforms (e.g., Oracle, Databricks). Integrating data architectures will likely require creating systems that can work between CSPs and platforms, because some CDC programs such as the National Syndromic Surveillance Program use AWS.

CDC is also promoting greater adoption of electronic case reporting (eCR) as part of its DMI programs. During the COVID-19 pandemic, CDC developed criteria for eCR, in which health care vendors directly provided eCR records to state health departments for rapid updates, and these state health departments collaborated with CDC for nationwide case monitoring. Applying a similar system to future pandemics and other public health emergencies will enable more rapid responses and resource allocation.

Public health reporting would also benefit from a central secure environment similar to the Trusted Exchange Framework and Common Agreement (TEFCA), which enables different health care providers (e.g., pharmacies, physicians) to communicate electronically. A centralized public health reporting network requires participation and trust by state health departments. To create that sense of trust and security, CDC is currently examining which TEFCA standards can be applied for a similar central public health reporting framework.

NCHS DMI Project Updates

James Craver, M.A.A., Deputy Director for Management and Operations, NCHS

NCHS currently has multiple DMI projects that focus on improving different aspects of data management and analyses, including data sourcing, production, and access. The projects described below highlight different aspects of CDC's DMI activities.

NVSS Data Modernization

Paul Sutton, Ph.D., Deputy Director, Division of Vital Statistics, NCHS

The NVSS modernization project has multiple goals:

- **Moving vital information faster** by working with jurisdictional partners and their data providers (i.e., upstream providers) to modernize vital statistic data exchange systems. NCHS also aims to improve interoperability through implementation of standards such as Fast Health Interoperability Resources (FHIR).
- **Improving data quality** by providing guidance, training, technology, and support to NCHS staff and jurisdictional partners.
- **Informing emergency response** through increasing usage of provisional data for public health and surveillance and response rather than relying solely on final data.
- **Enhancing data analysis methods** through data mining technologies such as machine learning (ML) algorithms, which can extract relevant data from text on vital records such as death certificates.

Enhanced Interoperability

Improving interoperability is a key goal of NVSS modernization because most current systems are siloed and require separate manual processes to pool and share vital statistics data. For example, mortality reporting currently begins with upstream data providers (e.g., funeral directors, coroners), many of whom have independent records systems, manually submitting mortality data through web forms to the state electronic death registration system. Once these data are in the state systems, data are then exchanged between those systems, the State and Territorial Exchange of Vital Events (STEVE) system, and NCHS. These data exchanges often use the Inter-Jurisdictional Exchange Format, but some systems use other data formats, requiring further manual data conversions. As a result, mortality records typically move between data systems in batches rather than as individual records are completed, and manual processes slow data movement and reporting.

Enhancing interoperability can provide multiple benefits, including (1) improved timeliness of both sending and receiving data through individual record-level data exchanges; (2) increased automation, which reduces amount of effort involved in data sharing and harmonization; and (3) improved data reliability and robustness using reliable data sharing architecture.

Adopting Health Level Seven (HL7) FHIR data standards can significantly improve electronic exchanges between different health information systems. FHIR uses standard data representations that enable reuse of health information by multiple parties for different purposes. The base standard covers 80 percent of health use cases, and FHIR standards can be extended to support the remaining 20 percent of use cases.

NCHS developed an implementation guide (IG) for the Vital Records Death Reporting (VRDR) FHIR standard, which maps data to current Inter-Jurisdictional Exchange data fields, enables submission of death records to NCHS, and returns coding responses to jurisdictional partners. NCHS also developed a vital records messaging FHIR IG, which provides a standard electronic system for messaging (e.g., submission and updates to death records) between NCHS and jurisdictional partners.

Greater usage of application programming interfaces (APIs) can also improve public health data interoperability by automating information exchange between systems. NCHS implemented a fully functional NVSS API that jurisdictional partners can test for data sharing and messaging, and the production NVSS API server will be implemented later this year. Jurisdictional partners can also download an open-source Reference Client API, which demonstrates implementation of the VRDR FHIR standard.

Vital Statistics Modernization Community of Practice

To facilitate vital statistics modernization, NCHS developed the Vital Statistics Modernization Community of Practice. This community built upon the previous Implementers' Community of 11 jurisdictions to provide greater sharing of resources for training, technical tools, and best practices for vital statistics modernization. The community is open to all 57 vital records jurisdictions and seeks to foster technical collaboration, promote greater adoption of common standards such as FHIR, and increase awareness of each participant's progress toward modernization goals. The community offers multiple resources, including monthly calls for members, a monthly newsletter, a common SharePoint knowledge management site, and quarterly virtual testing sessions on modernization approaches.

The community is also establishing a Steering Committee with representatives from NCHS, the National Association for Public Health Statistics and Information Systems (NAPHSIS), and jurisdictional partners. This Steering Committee will develop agendas for monthly calls and support jurisdictions in modernization efforts.

NCHS also supports the Community of Practice by hosting quarterly testing events. These events enable jurisdictional partners to test bidirectional interoperability between their systems, STEVE, and NCHS using the NVSS API. NCHS also uses these testing events to communicate changes and updates to VRDR FHIR standards and IGs.

Medicolegal Death Investigation (MDI) Reporting Modernization

The creation of the COMEC office provides an opportunity to further enhance interoperability for mortality reporting with coroners and medical examiners. NCHS is developing an MDI FHIR Implementation Guide to standardize MDI data reporting, and plans to develop an API to enable the exchange of data between the NCHS MDI case management system and jurisdictions' electronic death registration systems. NCHS along with four state jurisdictions initially tested the MFI FHIR standard in a connectathon in January, and based on this testing, the MDI FHIR standard met requirements to be published as a standard for trial use.

MedCoder System

NCHS is introducing MedCoder to replace its current system for coding cause of death, which was developed during the 1980s. MedCoder codes cause of death from literal text to International Classification of Disease (ICD)-10 codes using ML and natural language processing (NLP) approaches. MedCoder will provide multiple benefits, including record-level coding, a streamlined process for corrections, simplified maintenance, and ability to support FHIR-based record transmission.

NCHS tested the MedCoder system by reprocessing the entire NCHS 2020 mortality dataset through MedCoder and comparing results with the legacy coding system. This testing demonstrated consistency in coding between the legacy system and MedCoder, including records miscoded by the legacy system. Furthermore, MedCoder automatically coded approximately 85 percent of records, whereas the legacy coding system and the Mortality Medical Data System automatically coded only 70 percent of records. NCHS expects that further revisions and optimizations will enable MedCoder to automatically code greater than 90 percent of records.

MedCoder will be implemented in June 2022, with death coding system blackout beginning on June 6 that will last about 2 weeks. During this blackout period, 2022 data will be reprocessed through MedCoder so that all 2022 records will be coded using this system. Jurisdictions can still submit records during this blackout period, but the coding will not occur until the blackout period ends.

Improving Data Dissemination

Increasing usage of provisional data is one of the key objectives of NVSS modernization, including shortening timeframes for provisional data releases. NVSS modernization efforts have already reduced timeframes for releasing provisional drug overdose death data from 6 months to 2 months, and NCHS is focusing on reducing timeframes for other provisional data such as birth data and mortality data on the Wide-ranging Online Data for Epidemiologic Research (WONDER) system.

Discussion/Reaction by the Board

Dr. Das agreed with the importance of *bidirectional* communication for exchanging vital statistics data between NCHS and STLT jurisdictions, because historically federal agencies have emphasized communication to them rather than communication back to STLT partners. She also noted that many jurisdictional partners are not yet relying solely on death certificate text for cause of death coding because many jurisdictions are still training many medical certifiers on data quality for cause of death.

Dr. Sutton noted that NCHS previously returned coding data to STLT partners in batches, whereas adopting FHIR standards and real-time messaging will enable NCHS to return coding data within minutes

for approximately 85 percent of records that are coded automatically. NCHS has established a working group that collaborates with jurisdictions on improving data quality for cause of death. The Community of Practice will facilitate training and bidirectional communication to improve cause of death data quality and coding.

As noted above, testing of MedCoder on 2020 records showed that approximately 15 percent of records required manual coding. Approximately 5 percent of total 2020 records were automatically coded by MedCoder but then revised upon further review. For many of the other records that required manual coding, MedCoder could not automatically extract and code cause of death. Dr. Sutton expects additional refinements and optimization of MedCoder to further reduce the percentage of records requiring manual coding.

Division of Health Care Statistics (DHCS) Electronic Health Record (EHR) DMI Activities

Carol DeFrances, Ph.D., Acting Director, DHCS

Brian Ward, Ph.D., Chief of Ambulatory and Hospital Care Statistics Branch, DHCS

Askari Rizvi, M.S., M.B.A., P.M.P., Chief of Technical Services Branch, DHCS

Overview of National Health Care Surveys

DHCS administers National Health Care Surveys, which are a family of nationally representative provider surveys that cover a wide spectrum of care, including ambulatory, inpatient, outpatient, and long-term care. These surveys are based on patient encounters (e.g., hospital visit) and thus they provide estimates about encounters with health care providers rather than the population as a whole. DHCS' National Health Care Surveys also include provider surveys to gauge experiences of health care providers. The National Health Care Surveys are used to produce nationally representative statistics on health care to inform policy decisions and improve quality of care.

DHCS currently has four active surveys:

- National Ambulatory Medical Care Survey (NAMCS), which captures information from physician offices and community health centers.
- National Hospital Ambulatory Medical Care Survey (NHAMCS), which collects information on emergency department (ED) visits
- National Hospital Care Survey (NHCS), an all-electronic data collection that collects inpatient and ED data.
- National Post-Acute and Long-Term Care Study (NPALS), which obtains Centers for Medicare and Medicaid data for home health, hospice, nursing homes, inpatient rehab, and long-term care hospitals and does original data collection for adult day service centers and residential care communities.

NAMCS and NHCS are moving toward direct collection of data from EHRs. DHCS is also exploring using EHR data collection for NPALS, and it will collaborate with subject matter experts to assess availability of EHR data for long-term care sectors such as adult-based service centers and residential care communities.

Dr. DeFrances posed three questions to the BSC:

- Are there any additional supplemental data sources of hospital or ambulatory visits that may be worth DHCS investigating? Physician or provider groups/networks would be of particular interest.
- Are there any recommendations for managing Confidential Information Protection and Statistical Efficiency Act (CIPSEA)-protected data on the cloud?

- DHCS is currently migrating the Healthcare Electronic Health Record (HEHR) system to the EDAV cloud environment, which currently does not support SAS. Instead, EDAV supports DataBricks. Are any BSC members familiar or experienced with DataBricks? Are there any lessons learned or recommendations for migrating from SAS to DataBricks?

Moving to EHR Data Collection

DHCS leverages health care data and interoperability standards to facilitate movement toward data collection from EHRs. DHCS first used the HL7 Consolidated Document Architecture (CDA) implementation guide for EHR data collection. Moving forward, DHCS plans to transition to the HL7 Health Care Surveys FHIR Content IG, which is highly aligned with the United States Core Data for Interoperability (USCDI) standard as well as CDC’s MedMorph Reference Architecture, which is also used for eCR, Centralized Cancer Registry Reporting, and other public health and research use cases.

DHCS is planning a FHIR application pilot in 2022 and 2023. This pilot will involve four health care providers exchanging data with DHCS through the MedMorph application using the Health Care Surveys Content FHIR IG. Furthermore, this pilot will test the feasibility of building a data lake on EDAV and identify key elements for data governance that will enable information to be shared with other CDC CIOs.

DHCS is also leveraging health care interoperability regulations to advance EHR data collection. As part of the CMS Promoting Interoperability (PI) program, NCHS maintains the National Health Care Surveys Registry which is a public health reporting registry and currently has approximately 105,000 health care providers and 365 hospitals registered. The HL7 National Health Care Surveys CDA IG was also named as part of the ONC Cures Act Final Rule. DHCS is working with other parts of CDC to add other FHIR IGs to this Final Rule and the CMS PI program.

EHR Data Collection for NHCS and NAMCS

As noted above, NHCS collects ED and inpatient data from 608 non-federal and non-institutional hospitals with at least six staffed inpatient beds. Data are collected through administrative claims data as well as EHR data transmitted to NCHS or its data collection agent. NHCS collects personally-identifiable information (PII), enabling data linkages with other sources such as the National Death Index (NDI), CMS databases, and data from the Department of Housing and Urban Development (HUD).

DHCS is using DMI funding to enhance and modernize NHCS, including purchasing of additional data from the American College of Emergency Physicians and Premier Health. DHSC using NHCS data to model national estimates, examine the usability of synthetic data by the CDC Research Data Center (RDC), and use NLP algorithms to identify opioid-involved hospitalizations from EHR clinical notes.

DHCS is also examining EHR data collection for NAMCS, which collects data regarding ambulatory medical care services from 50 federally qualified health centers (FQHCs) and facilities classified as FQHC look-alikes (FQHC-LALs) by the Health Resources and Services Administration (HRSA). Similar to NHCS, the collection of PII enables data linkages with external data sources from CMS and HUD. DHCS is working with HRSA to recruit up to 50 FQHCs and FQHC-LALs to transmit EHR data using the CDA IG 1.2 version. DHCS is also providing financial incentives to health centers to implement or activate modules within their EHR systems that send customized reports to NCHS.

Building Infrastructure for EHR Data Collection

CDC’s HEHR system collects data from multiple data providers, including physician offices, hospitals, community health centers, and organizations such as the American College of Emergency Physicians and Westat. Data are currently stored in DHCS’s CIPSEA-compliant on-premise Structured Query Language (SQL) Server database. Users are then granted access to this database for analyses.

DHCS is using DMI funds to migrate HEHR from on-premise hosting to the cloud-based EDAV environment while still complying with CIPSEA and other privacy and security standards. DHCS is currently in Phase I of migration, in which the Registration, Annual Hospital Interview, and Annual Hospital Report portals are migrated to EDAV, and Phase I should be completed by the end of FY2022. During Phase II, HEHR data and the data transfer mechanisms will be migrated to EDAV. Migrating HEHR to EDAV will offer numerous advantages, including quicker processing times, streamlined configuration changes, and allowing other CDC departments to use HEHR's data transport mechanism.

Discussion/Reaction by the Board

Dr. Holan noted that the U.S. Census Bureau's Disclosure Avoidance System operated in a cloud-based environment that complied with Title 13 U.S. code, and it can provide lessons learned on cloud migration.

In addition to mortality reporting, DHCS is planning to modernize infrastructure for birth reporting based on lessons learned and infrastructure for mortality reporting. Initial efforts at modernizing mortality data have already highlighted necessary changes in traditional survey processes, and this information can be used to streamline the modernization of birth data reporting.

NCHS Virtual Data Enclave (VDE)

J. Neil Russell, Ph.D., NCHS RDC Director

VDE Need

Compared to a physical data center, hosting restricted data on a VDE presents an increased risk of disclosure, requiring CDC to assess whether the benefits of a VDE outweigh the disclosure risks. Based on this assessment, CDC initially limited the resources available to the VDE project to 10 virtual machines. However, a similar remote access system at the Census Bureau has more than 50 projects, raising concerns that VDE demand will exceed the current supply of 10 virtual machines. Furthermore, VDE's operational costs and researcher fees are currently unknown, raising questions about whether the VDE fee structure should differ from the RDC fee structure.

NCHS currently has four data access modes: (1) open data assets available for public use; (2) web-based data query system, which enables users to extract tabular data; (3) physical data enclaves such as the RDC and Federal Statistical Research Data Centers (FSRDC); and (4) VDE. VDE along with similar systems such as WONDER will likely play a key role in complying with the Evidence-Based Policymaking Act, which directs NCHS and other federal statistical agencies to expand access to data assets.

VDE Structure, Security, and Project Schedule

VDE allows approved researchers to remotely access and analyze restricted use CDC data rather than needing to travel to a physical data enclave. Researchers typically access VDE from an approved off-site location such as their researcher's worksite. Data are encrypted through a virtual private network (VPN) that requires two-factor authentication for login. NCHS VDE will be hosted on National Institute of Occupational Safety and Health (NIOSH) servers at CDC's Atlanta headquarters, which already have authority to operate (ATO) at the Moderate Impact level. NCHS data are segregated from other data on NIOSH servers to prevent comingling, and NIOSH employees and contractors overseeing these servers will be CIPSEA agents.

The VDE system allows NCHS to monitor and log user activity. Furthermore, the virtual machines prevent users from downloading, copying and pasting, or printing data. Certain Windows programs (e.g., Snipping Tool) that could enable researchers to capture and store restricted data have been disabled. Disclosure risk will be controlled through a data use agreement (DUA) between NCHS and the

researcher's employer, and this DUA will specify data protection requirements such as prohibitions on photographing or recording information displayed in VDE.

RDC received DMI funding for VDE in November 2020 and CDC approval to start the project in June 2021. The VDE project started in September 2021. RDC will begin to onboard pilot CDC projects in late summer 2022 and will onboard other projects in fall of this year.

VDE Benefits

VDE will provide multiple benefits, including the following:

- Addition of a tier of data access in compliance with the Evidence-Based Policymaking Act
- 24/7 data access that is not limited by pandemics, extreme weather, or civil unrest, which provides flexibility to researchers
- Decreased costs (e.g., reduced travel and background investigation costs) for researchers, which may expand the pool of researchers
- Use of current NIOSH servers with existing ATO, avoiding the need to obtain ATO for a new system

Discussion/Reaction by the Board

Dr. Holan inquired how the benefits of VDE differ from the current virtual RDC option that was created during the COVID-19 pandemic as well as why the VDE is not included under the Census Bureau FSRDC program. Dr. Russell responded that NCHS's current agreement with the Census Bureau does not allow NCHS to offer a virtual option. RDC has already started collaborating with NIOSH to establish the VDE on its servers, so establishing VDE at NIOSH will be quicker than amending NCHS's interagency agreement with the Census Bureau and then beginning VDE collaboration with the Census Bureau. Furthermore, hosting VDE exclusively through the Census Bureau would require users to undergo a background investigation for special sworn status, increasing the cost to researchers. Dr. Levy expressed concern that establishing VDE through NIOSH may not necessarily be quicker and encouraged Dr. Russell to reconsider establishing VDE through the FSRDC program.

RDC's current plan is to establish DUAs at the institutional level rather than with individual researchers. Dr. Malin noted that this approach will likely require amending the institution DUA every time a new user is added. Dr. Malin proposed the alternative approach of creating umbrella institutional DUAs and then investigator-specific DUAs that fall under that umbrella institutional DUA. Dr. Russell agreed that such an approach offers multiple advantages, including streamlining onboarding of multiple researchers from the same institution.

Dr. Peytchev expressed concern that the planned VDE would require reserving space for a single computer that can host the virtual machine and asked whether there could be different levels of data that are in the moderate environment that require different levels of security for more flexible modes of data access for the data users and increase the number of people who could be using the data. Dr. Russell replied that RDC is currently awaiting OMB's release of Section 3582 of the Evidence-Based Policymaking Act, which will provide more detail under different data security tiers. This additional detail will inform whether NCHS can use more flexible modes of data access.

NCHS Data Query and Visualization System: Development and Next Steps

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

Data Query System (DQS) Goals and Requirements

The NCHS DQS is designed to enable researchers and other data users to seamlessly connect to NCHS data assets with the least amount of required knowledge about the underlying individual data systems. For example, a researcher interested in diabetes statistics should be able to quickly access the DQS through the NCHS website to gain access to diabetes-related mortality and prevalence data from NCHS surveys such as NHCS and the National Health and Nutrition Examination Survey (NHANES). NCHS also seeks to enable users to draw explicit connections between data from different surveys and NCHS data sources as well as non-NCHS data sources. DQS is also designed to enable users to easily visualize these data to quickly identify trends or relationships.

Building a united NCHS DQS requires linkages across multiple NCHS and CDC data systems that cover a wide variety of health topics. Furthermore, establishing linkages across data sources requires a harmonized approach to health indicators or documentation when measures cannot be harmonized.

The *Health, United States* report, is a flagship publication that NCHS has provided to Congress since 1974. *Health, United States* uses both NCHS and non-NCHS data sources to create national-level estimates for health outcomes across different demographic groups. Because this report requires analysis of data across multiple data sources, including nongovernmental and Census Bureau data, NCHS identified the *Health, United States* report as an ideal source of pilot data for the DQS project.

DQS Advisory Group

NCHS formed an advisory group of subject matter experts on data queries, dashboard development, and other data management skills to develop a set of critical features for designing the DQS. This advisory group will also provide feedback on the pilot version of DQS built upon *Health, United States* tables and has developed a scope of work for the NCHS DQS contract. It will also continually identify further innovations to enhance DQS's capabilities.

Partnerships and Lessons Learned

During initial DQS development, NCHS partnered with other organizations both within and outside of CDC. NCHS is collaborating with the CDC Geospatial Research, Analysis, and Services Program (GRASP) team for system development, which has worked on projects such as the CDC COVID tracker, FluView, and U.S. Cancer Statistics. NCHS also partners with Tyler Technologies, which sells Socrata, a standard data platform used by NCHS to manage and store machine-readable data files and optimize searches. Socrata also provides multiple download formats and API endpoints for future integration with other data systems.

As described in the NCHS Director's Update presentation, NCHS has partnered with AcademyHealth for the undergraduate data challenge, which provided another opportunity to test DQS's capabilities. The winning team developed linkages by analyzing DQS files for differences in homicide rates between sex and age groups, and then linked these analyses with more detail on cause of death in WONDER. The contestants then used Tableau to visualize differences between gun-related deaths based on sex and age.

This contest provided an outlet for DQS products, including legacy print-friendly data tables converted into machine-readable data files. Contestants also highlighted areas for improvement such as improving the "cohesion" (i.e., harmonization) between data files and making methodological information more accessible.

Refined DQS Requirements and System Architecture

DQS requirements can be placed into three categories:

- **Timely:** Supporting future pandemic and public health emergency responses requires that DQS can quickly extract tabulations from existing channels, convert and validate these data for scientific accuracy, and then make these data publicly available. These tabulations should also be available in machine-readable formats to support modeling and forecasting.
- **Leverages Enterprise Resources:** Using existing NCHS resources and tools will provide a consistent user experience for CDC and NCHS users and enable interoperability with other NCHS and external data systems.
- **Customizable:** *Health, United States* users expressed strong interest in customizable and interactive data visualization tools (e.g., tables, charts, maps) as well as visualizations of data reliability (e.g., indicators of whether data are preliminary or final).

The DQS advisory group contributed to the proposed DQS system architecture, including data ingestion and data storage. DQS data come from aggregated, pre-tabulated data files that have been checked for scientific accuracy and reliability. Data files are ingested, and data are stored in the Socrata platform and federated to data.hhs.gov. The custom DQS application then accesses these data through Socrata APIs, and the DQS provides users with the capability to analyze and visualize data.

The DQS team is currently refining the DQS application based on these updated requirements and is planning to release the DQS pilot with data from *Health, United States* later in 2022. Following this pilot, NCHS will continue to develop and harmonize datasets for DQS and iteratively release updated DQS versions for public use.

Questions for BSC

Dr. Gindi posed three questions to BSC:

- What non-NCHS tabulated data sources would you recommend prioritizing for inclusion in the DQS?
- What query and visualization features would you recommend prioritizing for inclusion in the DQS?
- What user groups can we work with to promote use/feedback of early DQS products?

Discussion/Reaction by the Board

The DQS team continues to seek feedback on DQS features and usability, including from BSC members and other users outside of AcademyHealth. NCHS is using a similar approach focused on user experience and identifying user personas for redesigning the NCHS website. The DQS team will continue working with different stakeholders, including the CDC Communications Office, to further refine and optimize DQS.

Dr. Olson noted that DQS's data visualizations can be useful for educational purposes such as sociology and public health educational presentations, and she suggested promoting DQS on NCHS's social media accounts. Dr. Gindi agreed and described how the AcademyHealth contest illustrated the potential uses of DQS for education. Based on web surveys, many potential users are teachers and students, further underscoring the usefulness of DQS for different educational purposes. Dr. Lumpkin noted that the DQS team can also partner with other groups such as the CDC Foundation, de Beaumont Foundation, or the Robert Wood Johnson Foundation to facilitate similar contests or trainings for students in relevant fields such as public health and sociology.

Advancing DMI through Data Linkages

Lisa Mirel, M.S., Chief, Data Linkage Methodology and Analysis Branch, Division of Analysis and Epidemiology

Overview of NCHS Data Linkage Program

NCHS has a large amount of data across multiple surveys (e.g., NHCS, NHANES) that cover a wide variety of subjects such as health conditions, health behaviors, socioeconomic status, and health care access and utilization. Most household surveys are cross-sectional, and therefore researchers can identify exposures and risk factors. NCHS also has access to a large amount of administrative data such as housing assistance data from HUD and claims data from CMS databases. NCHS also geocodes the addresses of survey participants, which provides information on geographic variation and effects on health outcomes. The NCHS data linkage program focuses on linking these different data sources as an additional resource for public health analyses and decision-making.

DMI Activities for Data Linkage Program

NCHS recently linked NCHS survey data with Department of Veterans Affairs (VA) data (using DMI funding) and the CMS Transformed Medicaid Statistical Information System (T-MSIS), and these new data sources will be available to researchers during summer 2022. Linking NCHS survey data with VA administrative records will enable researchers to examine key policy questions such as health care utilization and health outcomes among veterans, both within and outside the VA system. Similarly, by linking NCHS survey and T-MSIS data, researchers can examine how health policy changes affect health outcomes and access among Medicaid recipients. NCHS is part of a community of practice within the HHS Office of Assistant Secretary for Planning and Evaluation (ASPE) to share information on the completeness of identifying variables in T-MSIS that were used for linkage.

NCHS is developing new methodologies such as identifying data linkages using custom-built SAS programs and ML technique algorithms. Based on feedback from the NCHS data governance board, the Data Linkage Program also compared its enhanced linkage algorithm to open-source linkage software. The program is currently exploring Match*Pro because it uses similar techniques in linkages (e.g., blocking groups) as those used by the current NCHS linkage algorithms.

Preliminary comparisons of linkage agreement between Match*Pro and the NCHS algorithm showed high concordance between the two techniques. Thus, Match*Pro could be used as an alternative linkage method. The Data Linkage Program continues to conduct comparison analyses, including assessing the impact of secondary analyses and resolving computing capacity issues for larger datasets.

NCHS is exploring Privacy Preserving Record Linkage (PPRL) techniques, which could expand linkage opportunities to new data sources while ensuring compliance with appropriate privacy and security requirements. The Data Linkage Program assessed linkages by the PPRL software Datavant with the existing NCHS algorithm and found that Datavant retained high precision and recall rates (above 93 percent) compared to the NCHS algorithm. The Data Linkage Program recently received additional funding from ASPE to continue PPRL assessments and research, particularly to expand analysis beyond NHCS data and explore other open-source PPRL software. The program is also assessing whether linkage quality differs between demographic groups. Ms. Mirel is participating in CDC's PPRL Community of Practice to gain additional insights and best practices.

Finally, the Data Linkage Program is developing methods for creating a public use synthetic data file that can enable more researchers to analyze linked data while protecting privacy. The program interviewed experts, including nine federal partners and three non-federal partners to identify key variables and potential research questions that could be answered with the synthetic data. Based on these interviews, the synthetic data files will focus on health equity and SDOH using the linked NCHS survey and Medicare and HUD. These synthetic data files should be available for researchers by summer 2023. Ms. Mirel asked the BSC to comment on whether the Data Linkage Program should also offer a validation service for users, which would provide greater certainty for researchers publishing results based on these synthetic data files.

Questions for BSC

Ms. Mirel posed three questions for the BSC:

- What are some innovative strategies we should consider for outreach about the new linked files?
- Are there concerns with synthesizing variables that have been brought in at the ZIP code level (e.g., percentage of ZIP code residents that are uninsured) for a person-level file?
- Do you have suggestions on how to set up a sustainable validation service for the synthetic data?

Discussion/Reaction by the Board

Dr. Lumpkin suggested contacting Lars Vilhuber at Cornell University regarding verification services and other best practices used for Cornell’s synthetic longitudinal business database. He also noted that generating synthetic geographic data from real data may still lead to unintended disclosures and recommended contacting the Census Bureau about approaches to avoiding disclosures. Dr. Malin described how similar PPRL validation research showed that reducing the risk of unintended disclosures from synthetic data can also reduce data fidelity and increase privacy costs. Dr. Peytchev recommended an approach of using diminished models for geographic variables to introduce additional variation (i.e., “noise”), thus reducing the risk of unintended disclosures.

Dr. Ohno-Machado agreed with Dr. Lumpkin’s concern about disclosures and noted that each additional linkage increases reidentification risk. She asked whether the Data Linkage Team has a protocol for mitigating reidentification risk and securing consent for data linkages. Also agreeing with the concern about reidentification risk, Ms. Mirel explained that the team is developing methods to quantify this risk. Regarding content, the Data Linkage Team only links data for people who have consented to data linkages, and data releases require approval by the disclosure review board.

The Data Linkage Program investigated software that performs homomorphic encrypted analysis for data linkages. However, this software had limited capabilities and could not generate models requested by many users.

Increasing Availability of Social Determinants of Health (SDOH) Data at HHS

Irma Arispe, Ph.D., Director, Division of Analysis and Epidemiology

Multiple recent federal initiatives have emphasized the importance of measuring and addressing SDOH, including the Evidence-Based Policymaking Act, an Executive Order on Advancing Health Equity, the White House Year of Evidence, the HHS SDOH Action Plan, and the ONC SDOH Activities.

HHS included evidence-building questions regarding SDOH in its FY 2023-2026 Evidence-Building Plan, such as how HHS programs and policies expand equitable health care access, promote healthy lifestyle behaviors, and expand access to mental health and substance use disorder treatments. This Plan also includes questions regarding improvements for data collection, analysis, and evaluation to reduce health disparities.

ASPE SDOH Resources and Reports

ASPE hosts multiple SDOH resources on its website, including the *Addressing Social Determinants of Health in Federal Programs* and *SDOH Action Plan – At a Glance*. These documents were released in March 2022 and are based on an SDOH framework developed under the Healthy People Initiative. The SDOH Action Plan outlines a three-pronged strategy that focuses on better data, improving health and social services connections, and “whole of government” collaborations.

ASPE also developed resources that examine SDOH among Medicare recipients based on the Congressional requirement for ASPE to study effects of socioeconomic status on health quality measures and resource usage under Medicare. In response, ASPE released the *Report to Congress: Social Risk Factors and Performance in Medicare's Value-Based Purchasing Program* in early 2020. This report highlighted three key questions regarding SDOH:

- Should some or all of Medicare's value-based purchasing programs account for social risk by adjusting measures and/or payment based on those measures?
- Should HHS routinely collect more extensive and detailed data on beneficiaries' social risk factors than is currently available?
- How can HHS achieve better outcomes for all Medicare beneficiaries by facilitating the ability of providers and communities to address social risk factors and integrate health and social services?

Released in March 2022, ASPE scoping review report on data elements for SDOH research highlighted the lack of standardized SDOH data elements, which limits SDOH research and data linkages. CMS also released a proposed rule in April 2022 for inpatient and long-term hospitals regarding health equity and SDOH measures. The proposed rule includes three measures related to health equity and also seeks stakeholder input regarding documentation of SDOH in inpatient claims data.

CDC SDOH Task Force

Based on the HHS SDOH Action Plan, CDC established an SDOH task force across CDC departments to share information and ideas regarding CDC's SDOH activities. The task force is organized around six categories: data and surveillance, evaluation and evidence building, partnerships and collaborations, community engagement, infrastructure and capacity, and policy and law. The task force identified multiple SDOH goals, including positioning CDC within the whole-of-government approach to SDOH, collecting and analyzing SDOH data, and providing resources for SDOH research. The SDOH task force also highlighted additional collaborations need to advance SDOH efforts such as the DMI alignment groups and the CDC SDOH Public Health Use Case Workgroup

NCHS SDOH-Related Activities

NCHS is conducting the following multiple activities to support SDOH analyses:

- Developing measures and analytic tools for understanding effects of SDOH on health
- Expanding SDOH data through data linkages and alternative data sources
- Developing model-based estimates of SDOH at state and local levels
- Developing new ICD-10-CM codes for SDOH
- Expanding analyses and dissemination of SDOH visualizations and analysis tools

NCHS is focused on standardizing SDOH measures and data standards as well as modernizing NCHS data systems such as surveys, NVSS, and other data assets. The diverse types of data collected by NCHS uniquely position it to support data analyses, including of how federal data infrastructure can be advanced and how SDOH impact clinical encounters.

Discussion/Reaction by the Board

As an additional resource, Dr. Williams suggested recent reports on SDOH by the National Academies of Sciences, Engineering, and Medicine, particularly one on capturing social and behavioral measures in electronic medical records. This publication reviews the status of scientific evidence that supports different SDOH measures and provides recommendations for improving how electronic medical records capture these measures. Dr. Snipp agreed and noted that the National Academies' Committee on National Statistics also routinely examines SDOH measures and can provide additional resources.

Workgroup Update: Assessing the Use of Commercial Web-based Survey Panels by NCHS

Andy Peytchev, Ph.D., Chair, Population Health Survey Planning, Methodology and Data Presentation (PHSPMDP) Workgroup

Background and Panel Meetings

NCHS tasked the PHSPMDP Workgroup with answering two questions: (1) Given current scientific knowledge, under what conditions would you recommend the use of online panel surveys for emerging or supplemental topics? And (2) What additional research and evaluation is needed to increase confidence in the fitness-for-use of estimates from online survey panels for these purposes? NCHS also suggested that the Workgroup consider multiple approaches, including supplementing online panel data with other survey modes, adding questions to regular NCHS surveys to link results with online panels, improving statistical approaches for integrating online panel data, and communicating data quality from online surveys.

Based on these questions, the PHSPMDP Workgroup decided to focus on incorporating external web-based panels rather than NCHS developing its own web-based panels. Web-based panels would be used to examine emerging topics (e.g., telemedicine usage during the COVID-19 pandemic) to supplement content already captured by existing NCHS surveys. NCHS would retain control over survey content and collected data, and these data would be disseminated through the NCHS Experimental Estimates program.

The PHSPMDP Workgroup held three panel meetings in early 2022 to gather information and discuss panel development and data dissemination. These panel meetings were attended by many experts from outside NCHS, including from the American Trends Panel, Gallup Panel, AmeriSpeak Panel, Census Bureau, and Kaiser Family Foundation.

Results from Workgroup and Panel Meetings

Based on panel and workgroup meetings, the PHSPMDP Workgroup identified key takeaways in response to the questions posed by NCHS.

Conditions for Using Online Panel Surveys (Question 1)

The workgroup agreed to avoid web-based panel surveys that duplicate information currently reported in NCHS surveys. Furthermore, estimates from online panels carry a greater risk of potential bias, and incorporating online surveys risks generating alternative estimates that differ from NCHS survey estimates.

Commercial web-based panels are well-suited for (1) timely estimates of new, emerging, and supplemental topics, including measuring changes over time; (2) repeated follow-ups with the same individuals; and (3) piloting new or altered measures prior to implementation within NCHS surveys. These uses require panel designs that remain stable over time. Furthermore, because NCHS does not directly manage commercial web-based surveys, selecting panels requires clarity on panel management decisions such as retiring or refreshing surveys.

The workgroup also agreed that NCHS should avoid using web-based survey panels that estimate measures related to data collection methods, willingness to participate, and some subgroup/subdomain estimates to avoid potential biases such as affluent minority and acculturation biases.

Additional Research Needed to Increase Confidence in Estimates from Online Survey Panels (Question 2)

Workgroup members suggested augmenting web-based survey panels with other samples or administrative data to address potential undercoverage and nonresponses to web-based surveys. Web-

based surveys can also be validated through evaluation studies such as a short omnibus survey with NCHS-relevant measures or a multi-panel study with additional cross-sectional arms for comparison.

Other methods for increasing confidence include evaluating different adjustment methods to account for potential sources of biases and including sociodemographic variables that can affect survey responsiveness such as technology usage and civic engagement. Finally, periodic evaluation of web-based panel surveys will ensure that survey structures function over time.

Vote on Report by the Board

The BSC voted unanimously to approve the report of the PHSPMDP Workgroup regarding the use web-based panel surveys.

Workgroup Update: Using NCHS Surveys to Measure Respondents' Perceptions of Discrimination

Kristen Olson, Ph.D., Chair, Workgroup to Consider and Assess Measures of Discrimination for Use in NCHS Surveys

Background and Panel Meetings

NCHS charged the workgroup with examining how best to add measures of discrimination measures to the NHIS, NHANES, and NSFG, with a particular focus on racism and other forms of discrimination based on nondominant social identities or positions. Incorporating discrimination measures into NCHS surveys can provide NCHS with capabilities to quantify and track the prevalence of discrimination as well as relationships between discrimination and key health indicators. Toward this end, the workgroup was established to gather information, conduct research on relevant issues, and prepare findings that can be used by the BSC to develop recommendations for NCHS.

NCHS tasked the workgroup with researching two key areas: (1) the feasibility and practicality of adding discrimination measures to NCHS surveys; and (2) additional research needed to select discrimination measures. The workgroup has held three meetings to date, interviewed key CDC stakeholders and experts, and reviewed existing literature on discrimination measures.

Preliminary Findings

Feasibility and Practicality of Adding Discrimination Measures (Question 1)

The workgroup affirmed that adding discrimination measures is feasible. Furthermore, NCHS can play an important role in research and surveillance of discrimination given its regular collection of nationally representative data. Data on discrimination from NCHS surveys can enable local communities to compare discrimination estimates from community studies with national benchmark estimates.

The workgroup agreed to focus both on overall experiences with discrimination and within health care contexts because capturing only discrimination in health care settings may miss key psychosocial factors that affect health outcomes. Furthermore, with similar discrimination measures (e.g., a core set of identical questions) across NCHS surveys, researchers can integrate discrimination measures across NCHS surveys.

Additional Research Needed for Selecting Discrimination Measures

Researchers have developed multiple major scales for discrimination that are applicable for general population surveys across racial and ethnic groups. However, these measures have not been explicitly connected in the same survey, limiting the understanding of these measures' relationships to health outcomes. Furthermore, past surveys have used different instructions and introduction questions for discrimination measures, so additional research is needed to understand how these instructions and introduction questions affect responses to discrimination measures.

Research on race- and gender-related discrimination attributes demonstrate differential responses based on the race and gender of the interviewer, so similar interviewer-related biases will likely occur for discrimination measures in NCHS surveys. Thus, additional research should explore this interviewer-related bias and its effects on variance in discrimination measures, as well how question order affects discrimination measures, whether existing measures capture the full severity of experiences with discrimination (e.g., interactions with law enforcement), and whether subgroups differ in how they interpret discrimination-related questions.

Future Information-Gathering

The workgroup has scheduled additional meetings with subject matter experts to assess multiple questions, including the following:

- What theoretical concepts related to which discrimination or measures are the strongest and most consistent predictors of health outcomes?
- Are some concepts and measures more useful for predicting specific health outcomes (e.g., chronic conditions, health behaviors)?
- What concepts or measures explain variance above and beyond information traditionally collected in NCHS surveys?
- Are certain concepts or measures “better” (stronger, more reliable, more likely to be endorsed, easier to be answered) for certain subgroups?
- How best can NCHS surveys contribute to future research and/or surveillance efforts related to racism and discrimination?

Dr. Olson also posed two questions to the BSC: (1) Which domains of methodological work should begin first? and (2) Are there additional criteria or evaluation questions that the workgroup should explore with subject matter experts?

Discussion/Reaction by the Board

Dr. Williams proposed examining linkages between measures of discrimination and SDOH, particularly psychosocial stressors. Recent research on vulnerable populations and people of color demonstrate that they experience a broad range of psychosocial stressors along with discrimination, suggesting potential linkages.

Dr. Brodie expressed concern about the large number of questions in upcoming subject matter expert interviews and encouraged the workgroup to identify the basic questions needed to begin collecting discrimination data and then iteratively improve measures from there. Dr. Olson agreed and noted that the subject matter expert interviews may help the workgroup prioritize research questions for implementing discrimination measures.

Dr. Peytchev noted that the three NCHS surveys (NHIS, NHANES, and NSFG) identified for potential inclusion of discrimination measures use different survey modes, which may impact responses to discrimination measures. He also suggested focusing on more immediate applications of these measures for NCHS surveys rather than large-scale implications for public health infrastructure. Dr. Williams responded that other large epidemiologic surveys have included measures related to discrimination to identify risk factors, and NCHS surveys should be able to include similar measures to identify risk factors and impacts on health outcomes.

Public Comment

Members of the public were invited to comment. The BSC did not receive any public comments.

BSC Wrap-up and Future Plans

Drs. Lumpkin and Moyer thanked attendees and speakers for their participation in today’s meeting, as well as emphasized their appreciation for efforts of the workgroups examining the feasibility of commercial web-based surveys and the inclusion of discrimination measures in NCHS surveys.

Dr. Moyer encouraged participants to suggest topics for future BSC meetings. Dr. Moyer also emphasized the importance of ongoing DMI activities across NCHS.

The meeting was adjourned at 4:42 p.m. ET.

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

/s/

October 4, 2022

John R. Lumpkin, M.D., M.P.H.
Chair, BSC

DATE