



Impact of the COVID-19 Pandemic on Major HHS Data Systems

KEY POINTS

- Major data systems that HHS relies upon for critical health and human services information have been disrupted by the COVID-19 pandemic.
- In-person data collections experienced an immediate disruption of operations while alternative modes of data collection were sought. This resulted in delays in field operations for some surveys and discontinuation for others.
- New content was added to both administrative and survey data systems to obtain pandemic-related information, including access to COVID-19-related care and the use of telemedicine.
- Some survey systems are responding with assessments of data quality and analytic guidelines on the use of data as affected systems will likely face issues of comparability for data collected during the pandemic versus in previous and future years.
- A number of major HHS data systems did not experience substantial disruption due to the pandemic. This was generally the case for administrative systems that did not collect information in-person and for surveys that collect information through telephone interviews.

BACKGROUND

Major data systems that HHS relies upon for critical health and human services information have been disrupted by the COVID-19 pandemic. The national response included unprecedented stay-at-home orders, guidance on social distancing, conversion to telework, school closures, and other dramatic societal disruptions, which complicated the collection and reporting of many kinds of health and human services information. Some data collection activities have been interrupted, others delayed, and others on hold indefinitely, while, at the same time, recent Executive Orders¹ highlight the critical need for information on vulnerable and underserved populations and the urgency of such information to ensure equity in the pandemic response. This brief identifies the impact of the pandemic on data systems used for key health metrics and measurement of progress and performance of Departmental priorities. It then proposes options and considerations for obtaining critical or priority data that may not be available from the usual HHS systems.

¹ *Ensuring an Equitable Pandemic Response and Recovery; Ensuring a Data-Driven Response to COVID-19 and Future High-Consequence Public Health Threats; and Advancing Racial Equity and Support for Underserved Communities Through the Federal Government*

The Department's core data collection systems provide the foundation for HHS' capacity to monitor the nation's health and well-being and track the functioning of health care and human services. These sources collect a wide range of information on health, health care, and well-being, including health status and behaviors, health care access, health care utilization and expenditures, insurance coverage and financing, functional status, as well as demographic data including race and ethnicity, employment, education risk-factors, and health conditions, quality of care, and well-being. HHS data sources are also broadly used by researchers, industry, and the public.

These data generally fall in two major categories: statistical surveys and administrative records. Surveys conducted by HHS collect information for statistical purposes, in order to describe, estimate, or analyze the characteristics of groups by collecting information from a representative sample, without identifying the individuals or organizations from which data are collected. These include both population health surveys, most often collected from households, and facility surveys, conducted in hospitals, physician offices, and other locations where health and human services are delivered. The Department's statistical surveys provide information on a broad range of health topics, behaviors, and characteristics related to health and health services in the US population. These data are widely used to monitor trends in health service utilization, insurance coverage, health behaviors, conditions and disability and to track progress toward achieving national health objectives. The data are also used by the public health research community for epidemiologic and policy analysis; such as characterizing those with various health conditions, determining barriers to accessing and using appropriate health care, and evaluating Federal and state health programs.

Administrative data generally describes information collected from individuals or entities, often in identifiable form, related to administrative, regulatory, law enforcement, adjudication, or other purposes that affect the rights, privileges, or benefits of a particular individual. These data are collected by HHS for programmatic, regulatory, surveillance, or enforcement purposes, but increasingly serve as valuable sources of information and evidence to inform policy and decision-making. The data are from systems designed for the administration of participant benefits in social service and public health programs, such as Medicare and Medicaid claims, and the data are crucial for evaluating programs and program improvements.

APPROACH

HHS has a substantial number of data systems and activities that provide information for a range of Departmental purposes – too numerous to summarize here. Therefore, the focus of this brief is restricted to major data systems that are relied upon most frequently and routinely by the Department and external users for purposes of measurement, monitoring program performance and progress, and conducting evaluation and research on the health and well-being of the nation. These data sources and their analysis contribute critically to informed decision-making. A number of these data systems were identified from a *Guide to HHS Surveys*

and Data Resources;² others from *Healthy People 2030*,³ including the leading health indicators (LHI).⁴ *Healthy People* states clearly that it relies on high-quality data to measure progress toward meeting national objectives; consequently, it is critical to assess any impact that might affect the data and its potential use. A limited number of additional data sources were included, resulting in a cumulative review of ten administrative data systems, 18 surveys, and one research program (Table 1 in the Appendix includes a description of each source reviewed).

To determine the impact of the pandemic on the selected data systems, a review was conducted of each relevant agency or program's website for pertinent information, as well as materials available from the Office of Information and Regulatory Affairs (OIRA) in the Office of Management and Budget (OMB), which encompasses collections of information that are subject to the Paperwork Reduction Act (PRA).⁵ The PRA gives OMB authority over certain collections of information by federal agencies in order to ensure the greatest public benefit and maximize the utility of information that is collected. In addition, the PRA has certain transparency requirements to make information publicly available regarding the purpose and use of data collections as well as methods and instruments, and served as a valuable source of information on the HHS data systems reviewed here.⁶ In addition to reviewing publicly available information, semi-structured interviews on the pandemic's impacts were conducted with key HHS staff that manage several of these surveys and administrative systems.

IMPACT

The impact of the pandemic on data systems has been experienced in a number of ways and differed depending on the type of system (e.g., administrative versus survey) and mode of data collection (e.g., in-person interview versus third party reporting). Below we describe the pandemic's effect on population health surveys, facility surveys, administrative data, and research programs. A summary of changes by data source is provided in Table 2 in the Appendix.

Population Health Surveys: HHS surveys that collect information using in-person interviews experienced the greatest impact. The Behavioral Risk Factor Surveillance System (BRFSS), which is conducted entirely over the phone, and the National Survey of Family Growth (NSFG), which was out of the field in 2020, were not notably

² ASPE, *Guide to HHS Surveys and Data Resources*, 2012. <https://aspe.hhs.gov/basic-report/guide-hhs-surveys-and-data-resources>. This is a compilation of information on the major data collection systems sponsored by HHS that are used to monitor the health of the population and the functioning of the health care system.

³ *Healthy People 2030*, <https://health.gov/healthypeople>. For Data Sources and Methods, see: <https://health.gov/healthypeople/objectives-and-data/data-sources-and-methods>. *Healthy People 2030* set data-driven national objectives with rigorous data standards for sources used to monitor progress toward national targets for improving health and well-being.

⁴ *Healthy People 2030, Leading Health Indicators, Data Requirements for Leading Health Indicators*. Available at: <https://health.gov/healthypeople/objectives-and-data/leading-health-indicators>. Data sources used to measure progress provide toward the Leading Health Indicators are comprised of reliable nationally representative data, are publicly available, and have no major methodological concerns.

⁵ <https://www.reginfo.gov/public/>

⁶ <https://www.reginfo.gov/public/>

impacted. All other population health surveys, including the following major surveys, appear to have been impacted:

- *National Survey on Drug Use and Health (NSDUH)*: Data collection for NSDUH is collected through personal interviews in respondents' homes. Data collection was halted in March 2020 due to restrictions on public interaction resulting from the COVID-19 pandemic. The program had planned to resume in-person data collection in October 2020, incorporating more versatile data collection methods that would include some in-person interviews as well as telephone and online interviews. Plans for new methods and COVID-19 related content were submitted to OMB for review in late September 2020; in May 2021, OMB approved the resumption of data collection.
- *National Health and Nutrition Examination Survey (NHANES)*: NHANES collects data through personal interviews in the home and also collects physical and laboratory data in mobile examination centers. In March of 2020, for the first time since continuous data collection began in 1999, NHANES was paused due to COVID-19. As a result, the 2019-2020 cycle will not be completed, and a 2-year nationally representative dataset will not be released. In response to the pandemic, NHANES proposed changes including modifications to the sample design, questionnaires, exam components, laboratory content, outreach materials, data collection modes, and operational procedures. Changes were submitted to OMB for review in late January 2021 and approval was given to resume data collection in April 2021; however, the program has noted that as a result of these changes, the next data collection cycle may not be comparable with other cycles. The survey was approved consistent with the understanding that National Center for Health Statistics (NCHS) will: a) conduct analyses to communicate the potential implications of these COVID-19 -related design and protocol changes for trend analysis, particularly subpopulation specific analyses and b) explore ways to use its monitoring trailers to better capture experiences in disadvantaged communities or subpopulations.
- *Medical Expenditure Panel Survey-Household Component (MEPS HC)*: Data are collected by personal interviews in the household for the MEPS HC. The survey moved to telephone interviewing in March 2020, for all panels and rounds currently in the field with increased use of the web to facilitate respondent reporting, and the survey plans to resume some face-to-face interviewing in the fall of 2021. In addition, some survey panels will be extended, in part to supplement the sample size, increase precision, and provide longitudinal capabilities to assess health care utilization and costs that will include the period of the pandemic and post-pandemic.
- *Medicare Current Beneficiary Survey (MCBS)*: A number of changes were implemented for the MCBS to ensure the health and safety of both respondents and field interviewers while continuing data collection. In March 2020, CMS paused in-person data collection in both community and facility settings, but the survey pivoted quickly to conduct interviews by telephone. The MCBS found some information more challenging to collect from Medicare beneficiaries over the phone, such as benefit and payment information that is usually gathered by trained interviewers in-person. A supplemental survey was added in summer and fall that consists of a phone interview, and a public use file provides data on the experiences of Medicare beneficiaries during the COVID-19 pandemic.

- *Health and Retirement Study (HRS)*: The HRS core or baseline interview is conducted in the home, with follow-up interviews conducted by phone or in person. The HRS halted the in-home collection of information and moved to telephone interviews. A supplementary 2021 COVID-19 Mailout Survey is currently in the field through August 2021.
- *Surveys of youth that collect information in classroom settings*:, National Youth Tobacco Survey (NYTS) and Monitoring the Future (MTF) had to discontinue field operations early, resulting in substantially smaller sample sizes. The Youth Risk Behavior Surveillance System (YRBSS) is conducted every other year and was not in the field in 2020. Data collection for the YRBSS in 2021 will be delayed from Spring to Fall.

Facility Surveys:

Facility surveys, such as the National Hospital Ambulatory Medical Care Survey (NAMCS), which typically collect information on-site from hospital records, had to instead utilize remote abstraction to access this information. Surveys that utilize visits to provider offices, such as the National Ambulatory Medical Care Survey, are subject to many of the same constraints as in-person interview surveys and had to stop or delay in-person data collection. In some cases, these delays were also used to add COVID-related questions to the survey content, as discussed below.

Administrative Data:

For some administrative systems, the pandemic presented initial challenges related to reporting. This was particularly the case for information collected from clinicians regarding what codes to use in clinical settings, for telehealth services, and for the certification of death. On the other hand, a number of programs did not experience delays or interruptions in the collection of information, and this was more often the case for administrative data systems than for surveys where the collection of information did not involve in-person reporting.

Administrative claims systems implemented changes to assist those providing information: new diagnostic codes, new procedure codes, changes to benefits and coverage to permit flexibilities in providing telemedicine, and exceptions to provider network rules. Medicaid has experienced increases in enrollment during the pandemic, resulting in more applications, verifications, and submissions to process. CMS has noted that the increase in enrollment is driven by the public health emergency and the enactment of legislation containing a continuous enrollment (maintenance of effort) requirement. CMS received nearly 700 submissions from states to make program changes to respond to the public health emergency for Medicaid waivers, State Plan Amendments (SPAs), Basic Health Program (BHP) Blueprint revisions, and other flexibilities as of mid-December 2020.⁷ In January 2021, CMS launched a new platform for section 1135 waiver requests for providers, as Medicaid waivers have increased due to the pandemic, and the web platform is expected to expedite requests and review.

For mortality reporting, the NCHS was receiving questions in early 2020 about how COVID-19 deaths should be reported on death certificates. To minimize inaccurate reporting, NCHS released preliminary guidance on March 4, 2020. This was followed by a National Vital Statistics System Alert on the new International

Classification of Diseases, 10th Revision (ICD–10) code for COVID-19.⁷ The introduction of a new code for the medical certification of death necessitated guidance and instructions for how COVID-19 should be reported on death certificates, and resources were produced for these purposes.⁸ Because mortality surveillance provides essential information that guides public health decisions, particularly in a public health emergency, these resources provided crucial information for accurate and timely reporting.

Other impacts were unique to specific systems. For example, the Treatment Episode Data Set provides information reported from state agencies on admissions and discharges for substance abuse treatment. In-person treatment ceased in many facilities for a period of time early in the pandemic; however, facilities were quick to respond with options for providing telehealth services. These changes altered the type of information reported. Additionally, The Mental Health-Client Level Dataset was affected by restrictions on some non-essential care; as a result, the data will show fewer individuals being treated. In another example, there was an increase in fraudulent claims reported to the National Directory of New Hires, which tracks both employment and unemployment. These claims were likely incentivized by the increased unemployment benefits available during the pandemic. Increased unemployment and economic distress also increased the flow of administrative data for human services as requests for services increased.

The switch to telework for many public health and human services professionals created some challenges to data collection and aggregation; these challenges would have been most felt at state, county, and city health departments with fewer resources. Some data systems experienced minor delays related to the switch to telework, particularly for information submitted via paper, as opposed to electronic forms. Staff responsible for reporting data were in some cases reassigned to COVID-19-related work or faced furloughs at the local level, but any ramifications of these changes also appear to be minimal.

Research Program:

All Of Us, a research initiative led by NIH, is distinct from the other types of data described above, but is one of the most ambitious research projects ever undertaken by the Department and is expected to generate valuable population health data for decades to come. The objective of All of Us is to enroll one million participants and collect health-related data throughout each person's entire lifespan through abstraction from electronic health records, bio-samples, and periodic online surveys. All of Us experienced a disruption in collecting bio-samples and physical measurements during the pandemic, however, the study is long-term and longitudinal in nature and does not rely on multiple collections of bio-samples. As a result, missed bio-samples can be collected at a later time and the pandemic has not meaningfully disrupted the program.

⁷ National Vital Statistics System, COVID-19 Alert No. 2, *New ICD code introduced for COVID-19 deaths*, March 24, 2020. <https://www.cdc.gov/nchs/data/nvss/coronavirus/Alert-2-New-ICD-code-introduced-for-COVID-19-deaths.pdf>

⁸ National Center for Health Statistics. *Guidance for certifying deaths due to COVID–19*. Hyattsville, MD. 2020. See: <https://www.cdc.gov/nchs/data/nvss/vsrg/vsrg03-508.pdf>

NEW CONTENT

A number of HHS data systems added new content to assess impacts and changes related to the pandemic, ranging from substantial changes, such as new survey modules, to minimal adjustments to specific questions. A summary of new content for each data source that made changes is provided in Table 3 of the Appendix. Some of these changes were implemented rapidly, such as the inclusion of new diagnosis codes in administrative systems for COVID-19, while other changes did not take effect until several months after the start of the pandemic. Selected changes, such as incorporating new diagnosis and procedure codes for COVID-19 and related treatment and including COVID-19 on the list of Notifiable Diseases, will likely remain in perpetuity, whereas new COVID-19-specific survey questions may be limited in duration.

For data collected from facilities, new data elements focused on shortages of personal protective equipment (PPE), COVID-19 tests, and staff; the volume of COVID-19 patients tested and treated; and changes in procedures and facility infrastructure to accommodate changes in the volume and type of patients. Examples include the following:

- NAMCS added five new questions on shortages of Personal Protective Equipment, such as N-95 masks, and COVID-19 tests, presumptive COVID-19 cases that were turned away or referred elsewhere, health care workers testing positive for COVID-19, and the use of telemedicine.
- The National Hospital Ambulatory Medical Care Survey added questions similar to NAMCS and also included a new item on whether areas were created outside the emergency department entrance to screen for COVID-19.
- Six questions added to the National Hospital Care Survey were also similar to NAMCS with the additional component of the number of inpatient and emergency department visits in the past year that were related to presumptive or confirmed COVID-19.

Survey data systems focused on individuals and households added questions regarding access, barriers, and the receipt of COVID-19–related care, as well as social distancing and use of telemedicine. For example:

- MEPS HC added new survey questions on barriers to accessing care due to COVID-19.
- NHANES added content to the questionnaire on COVID-19 tests, hospital stays, and treatment, and added a COVID-19 lab test to the venipuncture/phlebotomy component.
- NSDUH added questions about the use of telemedicine during the pandemic.
- The National Health Interview Survey (NHIS) and NHANES added new questions in 2021 on COVID-19 vaccinations.
- CMS took advantage of the MCBS panel design to assess and understand the COVID-19 pandemic by planning a series of rapid response surveys as a supplement to the main MCBS. The MCBS released out-of-cycle topic-specific Public Use Files on how COVID-19 was impacting the Medicare population (MCBS COVID-19 Fall 2020 PUF).
- HRS added COVID-19-related questions to the 2020 core interview and to the psychosocial self-administered questionnaire, and a special midterm data release is now available. Topics in the module include effects of the pandemic on health, access to testing and medical care, work, finances, and caregiving. Saliva collection kits were sent to 2020 sample respondents to test for SARS-CoV-2

antibodies. In addition, a new Contextual Data Resource on state-level COVID-19 policies has been added as part of the HRS restricted data products.

Research program

- The All of Us research program added a longitudinal COVID-related online survey, which was fielded monthly in mid- and late-2020.

CONSIDERATIONS FOR STATISTICAL SURVEYS

Although COVID-19 has had a substantial impact on data collection for HHS surveys, the full extent of this impact is not yet known and documentation on changes that agencies have made in response to the pandemic is not widely available. The majority of HHS surveys have no information at all on their agency survey webpages with regard to COVID-19 and any impacts on data quality. Most of the information regarding any pandemic impact is found in documents submitted to OMB requesting to change the mode of data collection or to modify other procedures or survey content. Often this information is quite limited in detail other than to note that in-person data collection was interrupted and transitioned to telephone interviewing, where possible.

Assessments of data quality: Some survey programs have released assessments of how impacts on their data collection may affect the resulting data and estimates produced from the surveys, and these programs are developing exemplar reports or guidelines that provide technical information on the use of the data in light of the impact of the pandemic. For example, NCHS released a report on a preliminary evaluation of nonresponse bias of NHIS estimates due to the COVID-19 pandemic that found a substantial impact on data collection beginning in the second quarter of 2020. The standard mode of data collection for the NHIS is through in-person interviews, and the survey had to change the mode from in-person to telephone. This change, in combination with the associated difficulty of identifying telephone numbers for households in the sample, contributed to a substantial decline in the response rate from 60.1% to 41.1% from the first to second quarter in 2020.⁹ The program also found some important differences in the characteristics of respondents, who were more likely to be older, more educated, higher income, and homeowners. This led the program to propose alternative approaches to weight the data to account for differential response and provided the public with information for consideration when using the data.

A significant change in a trend may reflect a real change in a population characteristic, but numerous other factors might be related to changes in estimate measurement. Federal statistical agencies and recognized statistical units routinely conduct these types of analyses, but this practice has broad applications for the

⁹ Dahlhamer JM, Bramlett MD, Maitland A, Blumberg SJ. *Preliminary evaluation of nonresponse bias due to the COVID-19 pandemic on National Health Interview Survey estimates, April- June 2020*. Hyattsville, MD: National Center for Health Statistics. February 2021. Available from: <https://www.cdc.gov/nchs/data/nhis/earlyrelease/nonresponse202102-508>

Department's statistical surveys conducted during the pandemic.^{10,11} Therefore, it is incumbent upon HHS survey programs to conduct evaluations of their results to assess potential bias and data quality given the unique circumstances pertaining to data collection throughout the pandemic and to propose solutions (e.g., generate weights to address bias identified).

The implications of disruptions to data collection will vary depending on the purpose and use of annual (or more frequent) estimates for different applications. For example, estimates for long-term monitoring of key indicators of the nation's health may be less impacted if data collection in subsequent years returns to pre-pandemic modes and response rates; in these cases, 2020 will likely be an aberration, but otherwise trends will remain valid. It will still be important, however, to attempt to examine what differences are due to changes in methodology and what differences reflect true changes in health or behavior during the pandemic. On the other hand, data used for evaluation studies or to assess the impact of programs will face significant challenges in discerning whether estimates reflect program or policy effects, changes related to COVID-19, or differences arising from the way in which data were collected during the pandemic. These types of policy or program analyses may be unable to draw definitive conclusions about impacts due to a lack of data that are comparable to the baseline. This will especially be the case for short term research examining programs during 2020 or the direct effects of COVID, particularly when this depends upon individual datasets that have been affected by the pandemic. However, whenever possible, HHS surveys and data systems that have been affected by the pandemic should provide a description of those changes on their public facing websites.

OPTIONS TO ADDRESS PANDEMIC DISRUPTIONS IN DATA SYSTEMS

The various disruptions to HHS data systems described above have and will create gaps in data needed for different applications. Outlined below are several approaches to help assess and manage these gaps, which may be employed alone or in combination, depending on the unique needs of each use of data.

Rapid surveys: Rapid surveys that can be conducted by telephone or online provide short-term alternatives and offer the flexibility to generate information on a range of topics in the absence of established data collection systems. One example is the Household Pulse Survey, led by the Census Bureau, which started during the COVID-19 pandemic to capture the social and economic effects of the pandemic on American households. There are recognized data limitations to such surveys and they cannot replace established household surveys; however, they can serve to address information needs as an alternative when major systems are interrupted. Rapid surveys can provide data that highlight health and human services priorities for Departmental programs in the short-term. For example, analysis of data from the Pulse survey found that, in January 2021, about 4 in 10 adults in the U.S. reported symptoms of anxiety or depressive disorders, an increase from 1 in 10 adults reporting such symptoms in the first half of 2019. This information can be informative for programs providing access to mental health services and implications for at-risk groups, particularly when estimates from traditional sources of this information may be unavailable.

¹⁰ SAMHSA. An Overview of Trend Testing Methods and Applications in NSDUH and Other Studies. 2017.

<https://www.samhsa.gov/data/sites/default/files/NSDUH-TrendTestTask1-2015.pdf>

¹¹ Ingram DD, Malec DJ, Makuc DM, Kruszon-Moran D, Gindi RM, Albert M, et al. National Center for Health Statistics Guidelines for Analysis of Trends. National Center for Health Statistics. Vital Health Stat 2(179). 2018.

Innovative methods: Innovative techniques and methods applied to existing data can increase and highlight value. For example, innovative statistical methods were applied by NCHS to mortality data to generate provisional estimates for many causes of death, and these innovations provided the foundation for reporting provisional deaths prior to the release of final mortality statistics with improved accuracy. More recently, nowcasting methods are being applied to provisional mortality data to estimate trends in drug-overdose, suicide, and transportation-related deaths, accounting for reporting delays and potential improvements in timeliness.¹² One objective of this effort was to examine how deaths from these causes might have changed due to any reporting lags resulting from the pandemic compared to prior years. There was little evidence of decreased timeliness of reporting since the beginning of the pandemic. This approach allows for the more timely release of provisional deaths, while continuing evaluation will inform future adjustment and improvements.

Alternative sources: If the generally observed pattern is that administrative data systems experienced less disruptions during the pandemic than surveys, a productive approach may be to supplement or replace indicators typically estimated from survey data with estimates based on related administrative data. Not all information collected via surveys will have a relevant replacement available from administrative sources, but triangulation using indicators that do have an alternative data source and creative uses of proxy indicators may help to assess the degree of bias in survey-based estimates.

New data collection: In addition to the example of the Pulse survey, described above, the COVID-19 pandemic opened other new avenues of data collection, which may also provide valuable alternatives to data collection that was interrupted or otherwise affected. For example, in mid-2020, HHS began requiring all hospitals in the U.S. to report several indicators on a daily basis through its Teletracking system. Daily patient volumes and stocks of personal protective equipment, as reported to Teletracking, may provide complementary information to measures routinely collected in hospital and facility surveys that were interrupted.

Weighting and other adjustments: Statistical surveys have also employed different approaches to adjust for disruptions to survey systems. For example, the NHIS currently uses weighting procedures that include steps that calibrate the survey's estimates to Census population estimates for certain characteristics. With the disruption in data collection processes and procedures, NHIS evaluated nonresponse bias due to the pandemic on survey estimates. This evaluation led to the inclusion of housing tenure in the weighting procedure, and while this does not entirely address differences, this inclusion appears to reduce nonresponse bias, although users of the data should be aware that additional bias in the estimates may still exist. And since NHANES' data collection was interrupted, the survey could not release nationally representative data and estimates for the 2019-2020 cycle. Instead, NHANES obtained nationally representative estimates by combining the previous data collection cycle with the partially completed 2019-2020 cycle and made adjustments to account for the incomplete data collection. The MEPS HC is collecting additional rounds of data to supplement the sample that experienced disruptions in data collection. This will enhance the sample size, increase precision, and

¹² Rossen LM, Hedegaard H, Warner M, Ahmad FB, Sutton PD. Early provisional estimates of drug overdose, suicide, and transportation-related deaths: Nowcasting methods to account for reporting lags. *Vital Statistics Rapid Release*; no 11. Hyattsville, MD: National Center for Health Statistics. February 2021. DOI: <https://doi.org/10.15620/cdc:101132>

provide information that will help assess data quality. The program will assess the additional rounds of data collection for potential bias and data quality to provide information to data users. These approaches provide users of the data with information that can help determine the fitness of the data for the intended use.

NEXT STEPS

Each HHS data system will need to assess the impact of the pandemic on data collection, analyze the impact on data values, provide guidance for using the data, and publicize these findings and recommendations to data users. Information on data quality should be disseminated in reports or technical notes, transparently providing users with assessments of the pandemic's impact and considerations for analysis and use. Agency assessments of the accuracy of the information and strengths and weaknesses of the data are important both for the use of data and for building trust and credibility with the public. The specific approaches for such assessment and adjustment, where needed, will likely vary by data system, but may include one or more of the options outlined above. It should be noted that many programs may not have the capacity to conduct these analyses internally as they have not been required to do so in the past.

The federal government has a responsibility to identify and address gaps in data that have emerged due to collection disruptions. OMB's Standards and Guidelines for Statistical Surveys recommends the development of methods to inform users regarding irregularities or special cases that must be taken into account in trend analysis.¹³ Survey programs should consider developing procedures to adjust data collected during the pandemic to improve comparability with data collected in other years. When collection systems have been affected by the pandemic, new data collections such as rapid surveys should take place or alternative data sources should be identified. This is particularly important for health conditions and social and economic risk factors that may have been affected by the pandemic, including mental health, substance abuse, obesity, employment, income, and social support. These conditions and characteristics will need to be prioritized for data collection.

CONCLUSION

The effect of COVID-19 on HHS data varies substantially across the Department's data systems and collection mechanisms. The ultimate ramifications for research and decision-making will depend both upon the reliance of each research question and decision on individual datasets that have been negatively affected during the public health crisis and upon the importance of a single year of data on those research and decisions. A number of major HHS data systems were not substantially impacted by the pandemic. This was generally the case for administrative systems that did not collect information in-person and for surveys that collect information through telephone interviews. In-person data collections, however, experienced an immediate

¹³ OMB Standards and Guidelines for Statistical Surveys. 2006.
https://www.cdc.gov/nchs/data/series/sr_02/sr02_179.pdf
https://nces.ed.gov/fcsm/pdf/OMB_Standards_Guidelines_Statistical_Surveys.pdf

impact on the ability to continue data collection while alternative modes were sought and implemented. The current situation regarding the difficulty of in-person data collection for key HHS surveys is clearly a significant challenge for those who rely upon information from these sources: researchers, health scientists, social scientists, clinicians, health-care workers, social workers, epidemiologists and decision-makers. However, there are options and alternatives, such as using rapid surveys and developing innovative methods, to address data needs and gaps resulting from interruptions in traditional systems.

Similar impacts on data systems have been experienced broadly across the federal government, particularly for federal statistical surveys, highlighting the need for assessments of data quality and accuracy. The Census Bureau has published information on the effect of the pandemic on survey response rates and the use of administrative data to evaluate non-response. Because of the impact of the pandemic, Census will not be providing the standard 1-year estimates from the American Community Survey (ACS), releasing instead a series of experimental estimates from the 1-year data. The Department of Labor likewise experienced decreases in response rates and in information reported to their household and establishments surveys that provide national-level estimates of employment in the U.S., resulting in some misclassification which the Bureau of Labor Statistics is evaluating. Both the ACS survey and the Department of labor's unemployment estimates are used by HHS and the impact of disruptions in those data collections must be considered by HHS analysts who should evaluate the data to determine fitness for use. These impacts on data collection have also been felt globally. The United Nations (UN) Department of Economic and Social Affairs concluded that "the pandemic caused a serious disruption to statistical operations across the world," while at the same time resulting in great demand for data to monitor and address the situation. By October, 75% of surveys conducted by National Statistical Offices were fully or partially pausing face-to-face data collection procedures, with "dire consequences" for key statistical programs.¹⁴ To address concerns regarding the impact of the pandemic, particularly for statistical surveys, transparent assessments of data quality are essential.

Recently issued Executive Orders and Presidential Memoranda highlight the need for data to measure and advance the Administration's commitment to equity for all Americans and to ensure that evidence-based decisions are guided by the best available science and data. These Executive Orders include *Ensuring an Equitable Pandemic Response and Recovery*, *Ensuring a Data-Driven Response to COVID-19*, *Future High-Consequence Public Health Threats*, and *Advancing Racial Equity and Support for Underserved Communities Through the Federal Government*. The Presidential Memorandum *Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking* affirms the Administration's commitment to the highest level of integrity in all aspects of scientific information and data. Together, these issuances specify requirements for data to address their core purposes, whether this is an equitable response to the pandemic, ensuring that there is enhanced data collection to drive decisions in the current pandemic as well as future public health threats, and that data collection is expanded to measure and advance actions to promote equity while adhering the principles of scientific integrity. The disruptions to data collection during the public health crisis will be particularly felt in understanding the consequences of COVID-19 across demographic groups and

¹⁴ United Nations, Department of Economic and Social Affairs Policy Brief #96, How the Data and Statistical Community Stepped Up to the New Challenges. March 26, 2021. <https://www.un.org/development/desa/dpad/publication/un-desapolicy-brief-96-covid-19-how-the-data-and-statistical-community-stepped-up-to-the-new-challenges/>

in addressing the effects of the pandemic on health and well-being resulting from foregone care, lifestyle changes, and mental health. An equitable, data driven response to the ongoing effects of the public health crisis will require additional data collections such as those outlined above.

APPENDIX

Table 1. Data Sources Reviewed

Data Source	Agency	Description	Website
<i>Administrative Data Systems</i>			
National Directory of New Hires (NDNH)	ACF	New hire, quarterly wage, and unemployment insurance information, reported by state agencies. Used to enforce child support payments and ensure appropriate payments by other federal agencies and programs.	https://www.acf.hhs.gov/css/training-technical-assistance/overview-national-directory-new-hires
Healthcare Cost and Utilization Project (HCUP)	AHRQ	National longitudinal patient-level health care database. Includes data collection from state data organizations, hospital associations, private data organizations, and the Federal government.	https://www.ahrq.gov/data/hcup/index.html
National Notifiable Diseases Surveillance System (NNDSS)	CDC	Nationwide collaboration that enables all levels of public health – local, state, territorial, federal, and international – to share notifiable disease-related health information.	https://wwwn.cdc.gov/nndss/
National Syndromic Surveillance Program (NSSP)	CDC	Collaboration among local, state, and federal levels of public health to improve nationwide situational awareness and enhance responsiveness to hazardous events and disease outbreaks.	https://www.cdc.gov/nssp/index.html
National Vital Statistics System (NVSS)	CDC/NCHS	Data on births, deaths, marriages, divorces, and fetal deaths from various jurisdictions legally responsible for the registration of vital events, provided for the nation’s official vital statistics.	https://www.cdc.gov/nchs/nvss/index.htm
Medicare Program Data	CMS	Information, including dates of service, diagnosis and procedure codes, and billing, submitted to CMS by healthcare providers for payment for services provided to Medicare Part A and/or Part B enrollees.	https://www.cms.gov/Medicare/Medicare
Medicaid Program Data	CMS	Beneficiary eligibility, beneficiary and provider enrollment, service utilization, claims, and expenditure information for Medicaid and the Children’s Health Insurance Program.	https://www.medicaid.gov/medicaid/data-systems/macbis/transformed-medicaid-statistical-information-system-t-msis/index.html

Uniform Data System (UDS)	HRSA	Data on patient demographics, services provided, clinical indicators, utilization rates, costs, and revenues, reported annually by HRSA-funded health center grantees.	https://bphc.hrsa.gov/datareporting/reporting/index.html
Mental Health Client-Level Data (MH-CLD)	SAMHSA	Client-level mental health data from states and territories, including demographics, clinical attributes, and outcomes routinely collected in monitoring individuals receiving mental health and support services.	https://www.samhsa.gov/data/data-we-collect/mh-cld-mental-health-client-level-data
Treatment Episode Data Sets (TEDS)	SAMHSA	Integrated system for reporting substance use and mental health client-level data for individuals receiving services, including the period of time in treatment, recovery process, and receipt of different types of services from different service providers.	https://www.samhsa.gov/data/data-we-collect/teds-treatment-episode-data-set
Surveys			
Medical Expenditure Panel Survey (MEPS)	AHRQ	Set of nationwide surveys of families and individuals, their medical providers, and employers. Data cover health care use and spending, insurance coverage, and accessibility of care.	https://www.meps.ahrq.gov/mepsweb/
Behavioral Risk Factor Surveillance System (BRFSS)	CDC	State-based telephone survey on health conditions and risk behaviors, ongoing annually since 1984.	https://www.cdc.gov/brfss/index.html
National Immunization Survey	CDC	A survey of children's immunization providers to monitor childhood immunization coverage.	https://www.cdc.gov/vaccines/imz-managers/nis/index.html
National Youth Tobacco Survey (NYTS)	CDC	Nationwide assessment of middle and high school youth's tobacco-related beliefs, attitudes, behaviors, and exposure to pro- and anti-tobacco influences	https://www.cdc.gov/tobacco/data_statistics/surveys/nyts/index.htm
Youth Risk Behavior Surveillance System (YRBSS)	CDC	Monitors six types of health-risk behaviors that contribute to the leading causes of death and disability among youth and adults.	https://www.cdc.gov/healthyyouth/data/yrbss/index.htm
National Ambulatory	CDC/NCHS	National survey on the provision and use of ambulatory medical care services, based on a sample of visits to non-federally	https://www.cdc.gov/nchs/ahcd/about_ahcd.htm

Medical Care Survey (NAMCS)		employed office-based physicians primarily engaged in direct patient care.	
National Health and Nutrition Examination Survey	CDC/NCHS	Survey on the health and nutritional status of adults and children nationwide, combining interviews and physical examinations.	https://www.cdc.gov/nchs/nhanes/index.htm
National Health Interview Survey (NHIS)	CDC/NCHS	National survey comprised of personal household interviews on broad range of health topics, ongoing since 1957.	https://www.cdc.gov/nchs/nhis/index.htm
National Hospital Ambulatory Medical Care Survey (NHAMCS)	CDC/NCHS	Information on the utilization and provision of ambulatory care services in hospital emergency and outpatient departments from a national sample of visits.	https://www.cdc.gov/nchs/ahcd/about_ahcd.htm
National Hospital Care Survey (NHCS)	CDC/NCHS	New survey that integrates inpatient data formerly collected by the NHDS with emergency department, outpatient, and ambulatory surgery data collected by the NHAMCS.	https://www.cdc.gov/nchs/nhcs/index.htm
National Post-acute and Long-term Care Study (NPALS)	CDC/NCHS	Started in 2012 as a survey of adult day service centers, assisted living and similar residential care communities, home health agencies, hospices, and nursing homes. Renamed to reflect the addition of more post-acute care sectors. Collected every two years.	https://www.cdc.gov/nchs/npals/index.htm
National Survey of Family Growth (NSFG)	CDC/NCHS	Survey on family life, marriage and divorce, pregnancy, infertility, contraception use, and men's and women's health. Mode is in-person interviews. Survey was out of the field in 2020. Interviewing is planned to resume in January 2022 through December 2029 pending funding.	https://www.cdc.gov/nchs/nsfg/index.htm
Medicare Current Beneficiary Survey (MCBS)	CMS	Survey responses and CMS administrative data for participants in a nationwide survey of a representative sample of Medicare beneficiaries. Includes information on health status and functioning and health care use and expenditures.	https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS

Monitoring the Future (MTF)	NIH	Survey of adolescent students on drug and alcohol use and related attitudes over their lifetime, in the past year, and in the past month.	https://www.drugabuse.gov/drug-topics/trends-statistics/monitoring-future
Health and Retirement Study (HRS)	NIH	Longitudinal panel survey of a representative sample of Americans over the age of 50, every two years, on changes in labor force participation and health transitions.	https://hrs.isr.umich.edu/about
National Mental Health Services Survey (N-MHSS)	SAMHSA	Annual census of facilities that provide mental health treatment services, including data on the location, characteristics, and utilization of organized mental health treatment services.	https://www.samhsa.gov/data/data-we-collect/n-mhss-national-mental-health-services-survey
National Survey of Substance Abuse Treatment Services (N-SSATS)	SAMHSA	Annual census of substance abuse treatment facilities, including location, organization, structure, services, and utilization.	https://www.samhsa.gov/data/data-we-collect/n-ssats-national-survey-substance-abuse-treatment-services
National Survey on Drug Use and Health (NSDUH)	SAMHSA	Annual survey that provides the primary source of information on the use of illicit drugs, alcohol, and tobacco among adults age 12 and older in the U.S.	https://nsduhweb.rti.org/respweb/homepage.cfm
Research Program			
All of Us	NIH	Longitudinal data collection effort, seeking to enroll over one million individuals, including participants from groups historically underrepresented in health research.	https://allofus.nih.gov/

Table 2. Data Sources with Identified Impacts of COVID-19 on Data Collection or Methodology

Data Source	Agency	Data Collection Methods	Impacts of COVID-19 on data collection or methodology
<i>Administrative Data Systems</i>			
National Directory of New Hires	ACF	Reported by employers to state agencies, which report to NDNH.	Changes in volume of data by type, due to reducing hiring and increased unemployment claims. Possibly some delay for information submitted using paper forms.
National Vital Statistics System	CDC/NCHS	Reported by state and jurisdiction vital registrars to NCHS	NCHS released updated guidance on mortality reporting around COVID-19 in March, April, and May of 2020.
Medicaid Program Data	CMS	Administrative data from states	Significant expansion of Medicaid enrollees
Mental Health Client-Level Data	SAMHSA	Reported by state agencies to SAMHSA.	17 out of 59 reporting jurisdictions experienced some delay in collecting or reporting information, particularly for the consumer satisfaction survey portion of MH-CLD.
Treatment Episode Data Sets	SAMHSA	Facilities report to state agencies, which report to SAMHSA.	15 out of 52 reporting jurisdictions experienced some delay in collecting or reporting information. Causes included staff reassignments to COVID-19 responsibilities, data entry staff furloughs, and delayed reporting by providers.
<i>Surveys</i>			
Medical Expenditure Panel Survey Household Component	AHRQ	Data collected using in-person interviews	Increased reliance on phone interviews and website interactions over in-person interviews
National Youth Tobacco Survey	CDC	Classroom surveys of youth	Data collected halted in March, reduced sample size. Switching to a web-based survey.

Youth Risk Behavior Surveillance System	CDC	Classroom surveys of youth	Data collection will be delayed from Spring 2021 to Fall 2021
National Health and Nutrition Examination Survey	CDC/NCHS	In-person interviews and physical examinations	Data collection paused in March 2020; May 2021 OMB approved to resume.
National Health Interview Survey	CDC/NCHS	In-person interviews conducted in the household	Temporary interruption in data collection March 2020. In-person data collection transitioned to phone interviews
Medicare Current Beneficiary Survey	CMS	Physical measures and interviews conducted as well as administrative data	Physical measures not being collected. In-person data collection transitioned to phone interviews.
Health and Retirement Study	NIH	Longitudinal surveys conducted in-person or by phone	Temporarily halted in-home data collection and moved to telephone
Monitoring the Future	NIH	Classroom surveys of middle and high school students	Data collection stopped prematurely for 2020, reduced sample size. Enough data collected to allow nationally representative sample through weighting.
National Survey on Drug Use and Health	SAMHSA	In-person interviews	Data collection paused in March 2020; May 2021 OMB approved to resume.
Research Program			
All of Us	NIH	Online Surveys, EHR sharing, DNA samples, physical measurements, and biosamples	Temporarily halted in-person bio-sample collection

Table 3. Data Sources with New Content Identified Related to COVID-19

Data Source	Agency	Content Added
<i>Administrative Data Systems</i>		
Uniform Data System	HRSA	Added data elements related to COVID-19 testing, diagnoses, and associated symptoms
Medicare Program Data	CMS	New diagnostic codes, new procedure codes, changes to benefits and coverage, and exceptions to provider network rules
<i>Surveys</i>		
Medical Expenditure Panel Survey	AHRQ	New survey questions related to barriers to accessing care due to COVID-19 added to the Household Component
National Notifiable Diseases Surveillance System	CDC	Added 60 new data elements related to COVID-19
National Ambulatory Medical Care Survey	CDC/NCHS	Added five new questions to survey covering shortages in personal protective equipment (PPE), shortages in COVID-19 tests, turned away or referred elsewhere presumptive positive COVID-19 patients, health care providers who tested positive for COVID-19 and telemedicine and telehealth
National Health and Nutrition Examination Survey	CDC/NCHS	New survey content to be added on COVID-19 tests, vaccinations, hospital stays and treatment as well as new lab tests (COVID-19 serology) for the venipuncture/phlebotomy component

National Health Interview Survey	CDC/NCHS	New questions added in quarter 3 2020 on COVID-19 related to testing, immunosuppression, access and delays to care, chronic conditions, caregiving and access to skilled caregivers, and social distancing at place of employment; Quarter 2 2021 added questions on COVID-19 vaccinations. Follow-back survey to be conducted with 2019 NHIS sample adults using the revised 2020 NHIS Q3 questionnaire. This will make it possible to compare the health of the sample adults in the year prior to the pandemic with their health after the start of the pandemic.
National Hospital Ambulatory Medical Care Survey	CDC/NCHS	Added new questions to cover shortages in personal protective equipment, whether they created areas outside the ED entrances to screen for COVID-19, if they turned away or referred elsewhere presumptive positive COVID-19 patients, and if they had any health care providers at their ED practice or center who tested positive for COVID-19
National Hospital Care Survey	CDC/NCHS	Added six new questions to survey covering shortages of COVID-19 tests, creation of COVID-19 screening areas outside the hospital, referrals for patients with confirmed or presumptive positive COVID-19 infection, clinical care providers testing positive for COVID-19, inpatient/emergency department (ED) visits for the year that were related to confirmed COVID-19, and inpatient/ED visits for the year that were related to presumptive positive COVID-19
Medicare Current Beneficiary Survey	CMS	Added supplemental phone interview of participants for experiences of Medicare beneficiaries during COVID-19
Health and Retirement Study	NIH	Added COVID-19-related questions to the 2020 core interview and to the psychosocial self-administered questionnaire. Topics in the module include effects of the pandemic on health, access to testing and medical care, work, finances, and caregiving. Saliva collection kits were sent to 2020 sample respondents to test for SARS-CoV-2 antibodies.
National Survey on Drug Use and Health	SAMHSA	Added questions relating to COVID-19 and the use of telemedicine
Research Program		
All of Us	NIH	Launched Covid-19 survey for research participants already in program