

Collecting Electronic Health Record Data for the National Ambulatory Medical Care Survey and the National Hospital Care Survey

by Carol J. DeFrances[†], Ph.D., and Denys T. Lau[†], Ph.D.

[†]Division of Health Care Statistics, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD, 20782

Abstract

The National Ambulatory Medical Care Survey (NAMCS) and the National Hospital Care Survey (NHCS) started transitioning into the electronic health record (EHR) realm with the 2016 data collection. With declining response rates, the National Center for Health Statistics (NCHS) moved NAMCS and NHCS to EHR data collection to reduce burden on sampled providers and hospitals to participate in the surveys. This paper first discusses key preparatory steps that NCHS took to move the first two of its health care provider surveys to EHRs, including conducting pilot tests, developing data standards, and leveraging incentives for participation. The paper then discusses implementation logistics and lessons learned during this transition, as well as the research capacity that the new NAMCS and NHCS EHR data will possess for use to better inform health care policies and actions to improve the health of the Nation.

Introduction

For more than 40 years, the National Center for Health Statistics (NCHS) has collected, analyzed, and disseminated data on health care utilization in the United States. These data have been used to examine the U.S. health care system and answer key questions of interest from health care policy makers, public health professionals, and health services researchers. To accomplish its mission, NCHS conducts the National Health Care Surveys, a family of surveys covering a wide spectrum of health care delivery settings from ambulatory and outpatient to hospital and long-term care providers. Specifically, the National Health Care Surveys include the following core data collections:

- National Ambulatory Medical Care Survey (NAMCS) collects data from office-based physicians and community health centers;
- National Hospital Ambulatory Medical Care Survey (NHAMCS) collects data from hospital emergency departments, outpatient departments, and ambulatory surgery locations;
- National Hospital Care Survey (NHCS) collects data from hospitals on inpatient discharge and visits to emergency departments and outpatient departments; and
- National Study of Long-Term Care Providers (NSLTCP) collects survey data from residential care facilities and adult day care centers and obtains administrative data from the Centers for Medicare & Medicaid Services (CMS) on care provided in nursing homes, home health agencies and hospices.

These surveys share certain design features in that each is designed to be nationally representative and collect data from health care providers. Because these surveys are of health care establishments, as opposed to households, the National Health Care Surveys provide

estimates about the universe of health care providers and the services rendered during the encounters with these providers.

The purpose of this paper is to discuss how NCHS is modernizing its data collection efforts from health care providers in the era of widespread adoption of electronic health record (EHR) systems. With the accelerated uptake of EHR systems in ambulatory and hospital settings, NAMCS and NHCS were the first of the National Health Care Surveys to transition to automated extraction of EHRs, starting with their 2016 data collections. The paper first describes the reasons why NAMCS and NHCS data collections moved to EHRs and then explains the preparatory steps that NCHS took to transition the surveys toward EHRs, including conducting pilot tests, developing data standards, and leveraging incentives for participation. The paper then discusses implementation stages and lessons learned during this transition, as well as new opportunities afforded by EHR data, including the increased research capacity that the forthcoming NAMCS and NHCS EHR data will possess to inform health care policies and actions to improve the health of the Nation.

How have data been historically collected for NAMCS and NHCS?

Historically, NCHS collected ambulatory care data for NAMCS and inpatient data for the predecessor of NHCS (called the National Hospital Discharge Survey) using manual abstraction of medical records, beginning with paper medical records onto paper forms. Manual abstraction of medical records was conducted by U.S. Bureau of the Census Field Representatives (FRs) on-site in physician offices and hospitals. FRs worked for several weeks in physician's offices and for a month or more in hospitals to abstract data from a limited sample of visits from these settings.

The abstracted data include information about the patient and encounter at sample visits. For example, patient's information include age, sex, race, insurance status, zip code, and medical conditions. Encounter information include the reason for visit, diagnoses, procedures, medications ordered and provided, diagnostic tests, and types of providers seen. Due to resource constraints such as time availability and finite space on the data collection tool, the maximum number of diagnoses, procedures, and medications captured per visit had to be limited. Furthermore, NAMCS and the National Hospital Discharge Survey did not historically collect personal identifiers such as name, address, and Social Security number.

In 2012, NAMCS was computerized so that data were abstracted using a computer-assisted tool instead of paper forms. After the 2010 data collection, the National Hospital Discharge Survey was discontinued and replaced by NHCS, which began in 2011 with all electronic data collection starting with Uniform Bill (UB)-04 administrative claims data. Then, in 2016, NCHS took a giant leap forward in transitioning to extraction of EHR data for NAMCS and NHCS to start modernizing the National Health Care Surveys into the electronic health record realm.

Why move NAMCS and NHCS to EHR data collection?

There are several reasons for transitioning the National Health Care Surveys to EHR data collection. The majority of office-based physicians and acute-care hospitals have adopted and

implemented EHR systems. Results from the 2015 National Electronic Health Record Survey showed that 88% of physicians had a basic EHR system¹. The American Hospital Association Annual Survey indicated that over 84% of hospitals had a basic EHR system in 2015².

Moreover, compared with onsite, manual abstraction of medical records, the burden on health care providers should be significantly less with electronic submission of data from the EHR. Providers will no longer need to give FRs physical space and computer access to the medical records for weeks or months at a time. Further, EHR data collection will eliminate abstractor error. Because data are transmitted electronically, errors introduced by abstractors manually entering data incorrectly into computer-assisted tools are avoided. Additionally, EHR data collection is more secure. Data are sent directly to NCHS or one of its designated agents thereby eliminating the need for physically transporting and securing portable laptop computers.

EHR data are more clinically detailed than manual abstraction because any clinical data in EHR can be captured without limitations as opposed to the manual abstraction process, which has limits on the number of diagnoses, labs, or medications that can be abstracted per visit. Finally, there is the potential for more clinical data in terms of volume per sampled provider. For NAMCS, EHR data can be extracted on all visits to a sampled physician during a reporting one-week period as opposed to manually abstracting data on a limited sample of visits in a reporting week. For NHCS, EHR data can be extracted on all visits to the sampled hospital's inpatient, emergency, and outpatient departments for an entire year, as opposed to abstracting data on a sample of inpatient discharges in a reporting one-month period.

What preparatory steps were taken for EHR data collection?

To prepare for EHR data collection, NCHS conducted pilot tests, developed data standards, and leveraged incentives for participation. First, NCHS conducted several pilot studies sponsored by Department of Health and Human Service Assistant Secretary for Planning and Evaluation. These pilot tests were conducted to assess whether the data extracted from EHRs would be comparable to the data abstracted traditionally from medical records. The findings indicated that clinical data extracted in the format of Continuity of Care Documents (CCDs), a standard EHRs can use to share summary information about a patient, coming directly from EHR systems were highly comparable to data manually abstracted from medical records for almost all variables collected in NAMCS. These findings gave assurance that extracted EHR data in CCD format will be comparable to much of the abstracted data traditionally collected in NAMCS³.

Because not all clinical data in different EHR systems are collected and stored the same way, NCHS needed to create a standardized format for submission of the National Health Care Survey data to enable automated extraction from the EHR or data repository. Data standards are the important technical underpinning to enable EHRs to share data effectively and efficiently among health care providers and between providers and public health agencies, like NCHS. In collaboration with the Office of the National Coordinator on Health Information Technology (ONC) and NCHS' Office of Classifications and Public Health Data Standards, the Health Level Seven International (HL7) Implementation Guide for Clinical Document Architecture (CDA®) Release 2: National Health Care Surveys, Release 1 - US Realm was developed to provide a standardized format and streamline the data collection for implementers to submit data to fulfill

requirements of the National Health Care Surveys. In 2015, the Implementation Guide was published as a draft standard for trial use and was described in the 2015 Interoperability Standards Advisory as the best available standard for clinical content and structure⁴.

NCHS also needed to find incentives to encourage providers to provide EHR data during a time when providers already have so many mandatory reporting requirements. Around the same time, the Centers for Medicare and Medicaid Services (CMS) EHR Incentive Program, also known as the “Meaningful Use (MU)” program, already started providing incentives for hospitals and providers to adopt and use certified EHR technology in ways that can improve patient care and meaningfully exchange healthcare information. Working with ONC, NCHS was able to leverage the MU program in 2015 as an incentive for participation by including the National Health Care Surveys as an option for eligible health care providers to meet the specialized registry public health reporting objectives, per the final rule on modifications to MU use in 2015 through 2017⁵. This means that eligible professionals, eligible hospitals, or critical access hospitals can submit data to the National Health Care Surveys to meet their public health objectives requirements to receive MU credit from CMS.

What were the implementation issues for EHR data collection?

During the 2016 transition year toward EHR, NCHS allowed for mixed modes of data collection. For NAMCS this included acceptance of traditionally abstracted data from some sampled providers and extracted EHR data from MU sampled providers. For NHCS, this involved acceptance of administrative claims data from some hospitals and EHR data from sampled MU hospitals. This enabled NCHS to do a proof of concept of collecting EHR data from providers and hospitals while still maintaining the original data collection mode for each survey.

New survey data collection procedures were designed and implemented. Instead of asking individual providers and hospitals for data directly, as had traditionally been done with abstracted data, NCHS had to approach provider and hospital networks as well as EHR vendors for EHR data because these data are stored centrally at the practice network level and require exchange of electronic health data from EHR systems that vendors build. Consequently, the number of questions about provider and practice characteristics traditionally asked in the NAMCS physician induction interview⁶ had to be decreased for those providers submitting EHR data because certain items could only be answered by sampled providers rather than by a network or EHR vendor staff person.

Per MU program requirements, NCHS created and implemented testing and validation procedures on the EHR data submitted by providers participating in NAMCS and hospitals in NHCS. EHR data were tested using a certification validation tool developed by the National Institute of Standards and Technology (NIST) to ascertain warnings and errors according to established HL7 CCD data standards. Results of the validation were shared with participating providers and hospitals so that they could address the errors and resubmit their EHR data. Only those providers and hospitals that have passed the testing and validation stage based on NCHS requirements were invited to submit full production data that would be used for official survey purposes.

Finally, NCHS started building an infrastructure and technical capacity to support EHR data collection, processing, and warehousing. A dedicated new secure platform had to be developed to handle the increased volume of data expected by collection of EHR data. Various transmission mechanisms such as DIRECT and secured file transfer program (SFTP) had to be explored. Hiring of new staff with SQL experience and data base specialists to manage the new data system was also needed.

How were 2016 NAMCS and NHCS EHR data collected?

NAMCS

As mentioned above, the 2016 NAMCS consisted of two modes of data collections: traditional medical record abstraction and EHR data extraction. The 2016 abstracted sample consisted of approximately 3,200 office-based physicians and 104 community health centers (CHCs). Office-based physicians and CHC providers had a sample of up to 30 patient visits abstracted for a preselected reporting week. Physician and practice characteristics were collected on physician specialty, office locations, number days worked and number of patients seen during the reporting week, cultural and linguistic practices, and use of EHR systems.

The 2016 EHR data collection was conducted in-house by NCHS staff. The EHR sample included about 500 office-based physicians and clinical data were collected for all of their patient visits during a preselected reporting week. No CHC providers were asked for EHR data. Because limited physician and practice characteristics were collected at the network level, new procedures had to be developed to collect physician induction information, such as physician specialty and number of visits seen in the reporting week, that was needed to weight the NAMCS data.

New procedures and infrastructure were built to develop and test software programs and create databases to store EHR data. A tracking system was constructed to document when letters were emailed, any email correspondence, and telephone calls. Written materials were produced to detail what providers were asked to provide, the data formats accepted, and how to use the CDC Secure Access Management Systems (SAMS) to send their test data and production data.

In June 2016, the first testing and validation letters were sent to physicians. The strategy formulated was to start with large physician groups since they had more technical staff to work with NCHS staff and then work down to smaller groups and individual physicians. Each group or provider was requested to have a technical call with NCHS staff within 30 days of being notified that they had been selected to participate in 2016 NAMCS.

At the technical calls, each participating provider was given several options to send data to NCHS including the HL7 Implementation Guide, a custom extract, or CCDs. For the 2016 NAMCS EHR data collection, all respondents chose to send CCDs.

NHCS

Unlike the NAMCS' samples of physicians and CHCs that are drawn anew every year, the NHCS sample of hospitals has been drawn and remains in the survey year after year until the hospital becomes out of scope or when NCHS decides it is time to select a new sample. In 2016,

the NHCS sample consisted of 581 non-federal, non-institutional hospitals with 6 or more staffed inpatient beds. Data collection was all electronic. Hospitals were asked to transmit UB-04 administrative claims data (billing data), Vizient (formerly University HealthSystem Consortium), or EHR data. All inpatient, emergency department, and outpatient department encounters were requested for a calendar year (January-December). Limited facility information such as total number of discharges per year, total births, total number of emergency department visits, and total number of outpatient departments visits are obtained which will be used to weight the NHCS data. Personally identifiable information such as name, address, and Social Security number is also collected for data linkage purposes. Due to low participation rates by sampled hospitals, NHCS has been unable to produce reliable national estimates of hospital care utilization since data collection began in 2011.

In 2016, among the 900 hospitals that registered with NCHS for MU credit, 98 were in the NHCS sample 581 of hospitals and only these 98 hospitals were targeted to send EHR data. The remaining 483 sampled hospitals could only provide claims or Vizient data. Testing and validation procedures were established for the NHCS data collection. Each hospital received a letter inviting them to testing and validation and indicating that they had 30 days to schedule a technical conference call. At the technical call, as done in NAMCS, hospitals were given three formats to transmit their EHR data: HL7 Implementation Guide, custom extract, or Clinical Document Architecture (C-CDA) (e.g., CCDs, Transition of Care, or Discharge Summary). For the 2016 NHCS EHR data collection, hospitals sent a combination of custom extract files and CCDA documents.

What are the challenges with EHR Data Collection?

NCHS is faced with technological, analytical and disclosure challenges with EHR transition. In terms of technical challenges, some hospitals and providers could not send 2016 EHR data to NCHS even with MU reporting requirements. These hospitals and providers corresponded with NCHS every 30 days to give updates on their process but had difficulty sending their data for various reasons including not being able to send their CCDs in batches. Interoperability remained a barrier. Even with the existing HL7 CCD standard, EHR vendors can tailor data elements according to the needs and workflow of specific provider networks. NCHS therefore had to invest substantial resources to develop custom programs to extract data elements for each vendor. Storing EHR data also remains a technical issue for NCHS. Due to the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) under which NCHS operates, NCHS decided not to use a cloud solution to store EHR data. Consequently, servers – an expensive alternative – are being procured to store the data in the interim.

NCHS also is grappling with analytical issues such as data integration and harmonization. For example, diagnoses codes in the CCDs obtained for both NAMCS and NHCS are in Systematized Nomenclature of Medicine (SNOMED) Clinical Terms and will need to be converted to the International Classification of Diseases and Related Health Problems, 10th revision, Clinical Modification (ICD-10-CM) as the standard coding template used by NCHS. In addition, given that EHR data come directly from clinical sources, the appropriate amount of EHR data processing, if any, will need to be determined for public health reporting and

benchmarking. Finally, disclosure review concerns are being examined to determine the best format and methods to release files while protecting patient confidentiality.

There were additional challenges in collecting 2016 NAMCS EHR data. The EHR data in CCDs were a reasonable match to meet NAMCS requirements but not a perfect match. Not all data elements needed for NAMCS were on the CCD documents. In addition, some physicians and physician groups could not retrieve EHR data for visits prior for a reporting week. No document submitted for testing and validation was error free according to the NIST validation tool. Some documents had a few errors, while others had hundreds of errors. For example, about 46% of the documents returned one or more errors regarding improperly formatted addresses. Over 70% of the documents had one or more errors on improperly formatted person names. About 10% of procedures and services were coded using implementation specific local codes rather than standard code sets such as Logical Observation Identifier Name and Code (LOINC) or Current Procedural Terminology (CPT).

For 2016 NHCS EHR data collection, there were additional challenges with custom extract and C-CDA documents. Custom extracts required a large time commitment by hospital staff to extract the data. It was often very hard to get hospital staff to provide data defining codebooks or reference files to look up hospital specific codes. There were also limitations on field size such as 2000 character limit for clinical notes.

C-CDA documents had many challenges as well. Some hospitals had to outsource work to create the connection with NHCS contractor and regenerate historical C-CDAs. Like NAMCS, no C-CDA documents received were error free. For almost half for of the hospitals submitting CCDs, it was difficult to definitively determine encounter setting (e.g., ED or OPD). Additionally, for ambulatory encounters, it was sometimes difficult to determine encounter start time and end time on CCDs.

What are lessons learned from EHR Data Collection?

Valuable lessons were learned from the 2016 NAMCS and NHCS EHR data collections. First, the MU program is a powerful incentive. Many physician groups and hospitals were eager to talk with NCHS staff and be actively engaged to send EHR data if asked. Second, NCHS gained valuable insights by conducting NAMCS EHR data collection in-house in 2016. NCHS staff worked first-hand with physician groups and hospitals and their technical staff and heard the challenges providers faced in sending EHR data. Moreover, NCHS staff had the opportunity to examine CCDs sent from different EHR systems and assessed their nuances. Third, NCHS discovered that C-CDA documents were more easily available to produce for NAMCS providers and NHCS hospitals because C-CDA documents were already programmed into EHR systems. Custom extracts required much more resources to produce by providers. Further, solo-practice physicians could participate in NAMCS with the help of their EHR vendors. Fourth, because of EHR vendor variations in capturing clinical data, NCHS learned that additional resources were required to develop custom coding to acquire common data elements. For example, diagnoses were captured in different structured and unstructured fields. Additionally, some EHR vendors used their own local coding set instead of CPT or LOINC codes, which are industry standards, for capturing procedures. Finally, NCHS found that the information extracted from CCDs

provided most of the data needed for NAMCS and NHCS. However, EHR data affords the opportunity for more data to be available that are not currently being collected for NAMCS and NHCS. Examples include allergies to medication, family history, social history and alcohol use.

What will be the increased research capabilities with EHR data?

Collection of EHR data provides a tremendous opportunity for the National Health Care Surveys. As the National Health Care Surveys are transitioning toward EHR data collection – first with NAMCS and NHCS, the surveys will have more clinical richness, depth, and increased research capabilities. All diagnoses, active problems, medications, as well as diagnostic imaging and lab tests ordered and their results listed in the EHR can be obtained. The data will provide a unique opportunity to study rare conditions and new procedures that have not been studied before with a large-scale database. For those surveys that collect personally identifiable information, such as patient name and address, data will be able to be linked across sampled health care settings and can provide information about readmissions or revisits. Data can also be linked to other data sources, such as the National Death Index, to allow the reporting of 30-, 60-, and 90-day mortality after the survey encounter. With new analytical capabilities comes the potential for new and exciting health services and outcomes research that generate insights to inform health care policies and actions to improve the health of the Nation.

References

1. NCHS, 2015 National Electronic Health Records Survey (NEHRS).
2. Office for the National Coordinator for Health Information Technology/American Hospital Association (AHA), AHA Annual Survey. 2015 Information Technology Supplement.
3. Gugerty B. NAMCS EHR Pilot 2: Overview and Select Findings. Presentation to the Division of Health Care Statistics, National Center for Health Statistics, August 15, 2016.
4. Office of the National Coordinator for Health Information Technology. 2015 Interoperability Standards Advisory. March 2015.
5. Medicare and Medicaid Programs; Electronic Health Record Incentive Program-Stage 3 and Modifications to Meaningful Use in 2015 Through 2017: A Rule by the Centers for Medicare & Medicaid Services on 10/16/2015. Federal Register: <https://www.federalregister.gov/documents/2015/10/16/2015-25595/medicare-and-medicaid-programs-electronic-health-record-incentive-program-stage-3-and-modifications>
6. See Physician Induction Card https://www.cdc.gov/nchs/data/ahcd/2016_NAMCS_Traditional_Physician_Induction_Interview_Sample_Card.pdf