

The Feasibility of Utilizing EHR Data to Supplant Survey Data Collection: A Case Study

Kailee Zingler

University of Chicago - Graham School MScMBI Student

Prashila Dullabh and David Liebovitz

Senior Fellow NORC and CMIO at UCMC

This case study assesses the feasibility of use of Electronic Health Records (EHR) for healthcare quality surveys, such as the National Hospital Care Survey (NHCS), in lieu of paper surveys. The NHCS became the first electronically collected government healthcare survey in 2011. It was founded by the Centers for Disease Control (CDC) with the goal of collecting reliable and nationally representative data on trends such as the opioid epidemic in America. For the purpose of this study, EHR data quality, collection, standardization, extraction, transmission, and secondary use was researched through a literature review. Key informant interviews were carried out for gathering expert opinions on the feasibility of the NHCS. Secondary use of EHR data was found to be considered important for the future of healthcare research. The methods of extracting the data from EHRs were not agreed upon across the literature or interviews. The NHCS recruitment currently falls short of collecting nationally representative data. The survey may benefit from taking part in more incentives programs, in order to off-set the costs of the EHR extraction methods required.

Introduction

The National Center for Health Statistics, within the Center for Disease Control (CDC), utilizes National Health Care Surveys (NHCS), which collect patient visit information from across clinics and hospitals (Brown, 2015). The National Hospital Care Survey (NHCS), within the NHCS, became the first government survey to be filled with electronically collected HIPAA compliant Private Health Information (PHI) (DeFrances, 2012). The data required for the NHCS, already collected in EHR systems during clinic visits, pulls patient information on demographics, insurance status, residential zip, medical conditions, smoking history, and personal identifiers, as well as, the following encounter data: reason for visit, diagnosis, procedures, medications, laboratory and diagnostic tests, and types of providers (DeFrances, 2012).

Meaningful Use, a Centers for Medicare and Medicaid Services (CMS) EHR Incentive program, provides incentives to support EHR adoption in clinic settings for improvement of patient outcomes (Blumenthal, 2010). This incentive program included Meaningful Use requirements such as automated reporting of quality measures, and public health reporting (Adler-Milstein, 2011). Meaningful Use has a public health reporting measure where physicians receive credit for submitting data for the NHCS. This allowed for the CDC to create and encourage use of their Health Level-7 (HL7) guide, now the best available standard for national health care surveys (CDC, 2018), (HL7, 2017), (DeFrances, 2017). HL7 is a set of internationally recognized and commonly used standards, or guides, for the transfer of clinical

information between different software and organizations (Dolin, 2001).

The information gathered for the NHCS allow for the possibility of providing national health statistics on opioid related incidents (Brown, 2018). This survey has the potential to track relevant trends like drug abuse or heart attack instances (DeFrances, 2018). The NHCS may also be able to link encounters across hospitals as well as to other data sources such as the National Death Index (NDI), not currently available to hospitals (DeFrances, 2017). The importance of this survey and surveys like this intend to be highlighted in their ability to track and link drug abuse encounters across hospitals as well as these instances of death to the NDI.

The goal for this case study was to understand the NHCS, the informatics topics surrounding it, the CDC goals for this survey, and how to ensure increased success in participation.

Methods

This case study explored, through 1) literature reviews and 2) interviews, the feasibility for the NHCS electronic data collection model as well as for hospitals to volunteer as participants. This study was conceived by Prashila Dullabh to serve the purpose of fulfilling the graduation requirements for completing a capstone project by the author.

The literature review was completed between February and October of 2018 through PubMed, Google Scholar, Science.gov, JSTOR, and the University of Chicago Library with the following combined search terms: EHR data, HL7, healthcare surveys, and the opioid epidemic. The various

search terms shown in Figure 1 below resulted in hundreds of articles, which were narrowed down twice, first to 35, and then to 24 articles. The articles with the most relevant titles were then paired down based on abstracts, which seemed to provide the most pertinent information on electronic survey data collection. Finally, the articles chosen had full text available, and were from peer-reviewed journals. Information collected from the literature review provided a foundation of knowledge on government surveys, healthcare regulations, EHR data collection and transfer, HL7 guides, and the opioid epidemic.

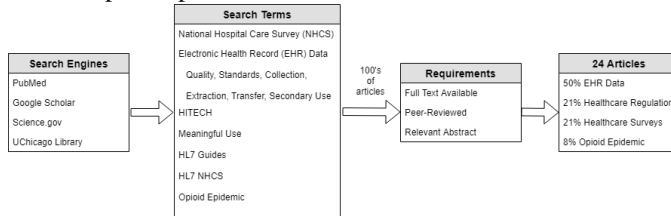


Figure 1: Literature Review schematic

Key informant discussions took place with experts on EHR systems introduced through capstone leadership as well as the point of contact for the NHCS at the CDC contacted through email. The key-informant discussions were structured interviews directed by questions referred to here on out as the vendor guide and shown in Table 1 below. The questions centered around the challenges with electronic surveys, specifically NHCS and other NCHS. We discussed hospital experiences in implementing specific HL7 implementation guides to gauge whether this would facilitate success for NHCS.

Key Informants	Interview Questions
Source 1 • NHCS expert at CDC	EHR Data • Secondary Use • Reporting • Extraction
Source 2 • Informatics expert at UChicago	Healthcare Surveys • Data Collection – Real Time • Burden
Source 3 • EHR reporting expert at UChicago	NHCS • Importance • Incentives/Recruitment
Source 4 • Telehealth entrepreneur	Survey Data Extraction • EHR Reporting • HL7

Table 1: Key Informants and Interview Questions Schematic

The key informants and their areas of expertise are broken down in Table 1 above. A key informant interview took place with Dr. David Liebovitz the Chief Medical Information Office of the University of Chicago Medicine on September 21, 2018. On September 25, 2018, an interview was conducted with Dr. Carol DeFrances Branch Chief, Ambulatory and Hospital Care Statistics Branch of the US Department of Health and Human Services. A key informant discussion was carried out on October 5, 2018

with John Moses, Director of Enterprise Architecture and New Technologies for University of Chicago Medicine.

The literature review provided initial information on EHR data secondary uses hurdles and possibilities. The key informant discussions highlighted the real world considerations in using EHR data in lieu of survey data collection.

Results

Literature Review

Figure 2 below displays the breakdown of how many papers were pulled per category of search during the literature review.

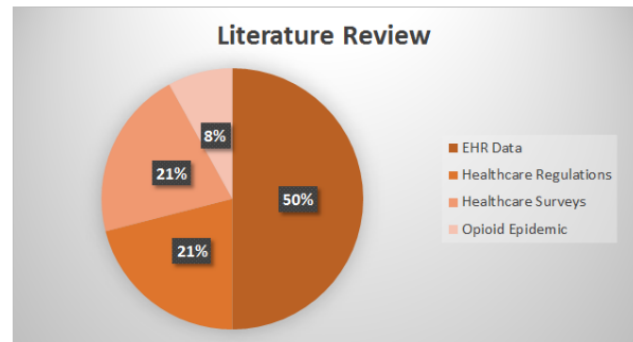


Figure 2: Literature Review papers pulled, percentage of category per total

National statistics on EHR adoption show that as of 2015, 96% of all hospitals had adopted certified EHRs (HealthIT.gov, 2017). Model based estimates reported that widespread EHR adoption and implementation have the potential for billions of savings annually in healthcare costs (Hillestad, 2005). Electronic data reporting allows for faster and more accurate quality measure reporting in healthcare which is important for reimbursement, ratings, and retaining patients (Weiner, 2012). Teams such as Rajeev, et al., have created their own internal systems for generating electronic public health care reports (Rajeev, 2010). In 2006, a panel of informaticians discussed the landscape, at that time, of utilizing EHR data for secondary uses; the consensus was that governments would need to play more active roles in the future of EHR data secondary use (Safran, 2007).

A study by Scholte, et al., compared paper surveys to EHR extracts, finding that changes in data collection methods between the survey and EHRs created the largest impact on result comparison, observing that the EHR collection had fewer missing data points on average than the surveys (Scholte, 2016). Wagaw, et al., completed a study in which they had patients complete self-administered surveys measuring seven continuous health behaviors, such as using a seat belt and eating fruits and vegetables as well as

eighteen categorical variables and found self-reporting issues in various categories (Wagaw, 2018). Self-reported health surveys and EHRs were compared over multiple diseases by Violan, et al., it was discovered that patients under-reported mental health disorders versus their doctor's; indicating a path in which electronic health information could serve a positive purpose as a source of truth (Violan, 2013). Botis, et al., conducted a review of survival status among a group of patients at Columbia Medical center, International Classification of Diseases, Clinical Modification (ICD-9-CM) codes to identify patients. The data queried was found to be incomplete, inconsistent, and the information reported was not always correctly recorded in a standard manner, leading to difficulty in secondary use of the EHR data (Botsis, 2010).

EHR data is either presented as structured coded data or unstructured narrative data (Chan, 2009). There is potential issue in the reliability and validity of EHR data. At this point in time, there is not a solution to entice healthcare staff to fill in all fields in EHRs in a uniform manner. There also does not appear to be standardization across EHR vendors in which fields are structured or not (Chan, 2009). A majority of the EHR data is not considered useful or structured sufficiently for data analysis. Without data standardization, natural language processing is required to process EHR data, which is complicated and time consuming (Murdoch, 2013). The HL7 Clinical Document Architecture (CDA) provides a set of standards for clinical documentation (Dolin, 2001). CDA v2 can transfer all aspects of the original clinical notes in HL7 messages and this data can exist outside of the messages (Dolin, 2006).

HL7 interfaces have become the industry standard for transfer of health information (Source 4). The v2 guides are expensive, commonly around five-thousand dollars per system upload, and usually require maintenance through the IT Department, an interface group, and a Physician Champion, as well as going through compliance checks (Source 4). Hospitals are expected to pay for and acquire HL7 guides and maintain them, for their individual compliance requirements (Source 4). EHR systems tend to have in house reporting workbenches with normalized data available for queries (Source 2). Providing extracts from EHR reporting workbenches would be cheaper, quicker, and easier (Source 3). The expert opinions on the subjects of EHR data extraction and transfer, as well as survey burden and NHCS importance are displayed in Figure 3 below.

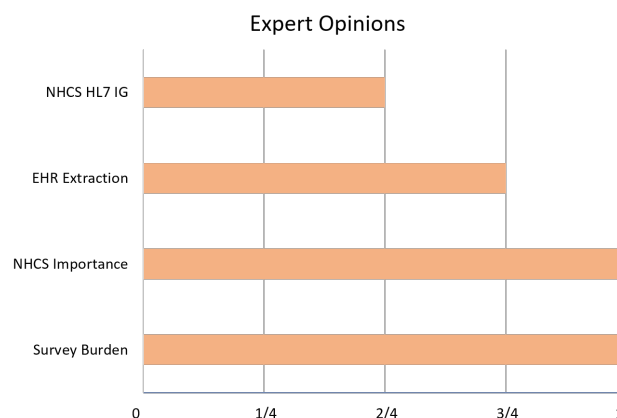


Figure 3: Key Informant opinions in favor

Meaningful Use created a public health reporting measure in which physicians would get credit for taking part in the NHCS (Source 1). The NHCS HL7 v2 implementation guide (IG) was required to be created in order for all of the NCHS to participate in Meaningful Use (Source 1). The IG also served the purpose of relieving burden on each participant to pull the requirements from EHR data themselves (Source 1). The HL7 IG allows for the CDC to receive all of the data in a standardized format (Source 1). Taking part in Meaningful Use was important for the survey to gain traction and offer greater incentive to participants beyond sharing the nationally representative and NDI data (Source 1).

Networks of academic medical centers, community hospitals, integrated health delivery networks, etc., utilize a data warehouse through the organization Vizient (Source 2). The members receive clinical quality measure information as well as supply chain improvement information (Source 2). Individual data warehouses and their reporting workbenches are retrospective, as opposed to the real-time reporting of HL7 messages, but still allow for researching important quality measures (Source 3). The NHCS has goals for sharing and conducting review of data collected in regards to the opioid epidemic (Source 1). This survey would like to follow drug related visits across hospitals and possibly help to prevent opioid overdose in the cases of drug diversion illustrated in Figure 4 below (Source 1). Notable issues encountered thus far with the NHCS have included recruitment, interoperability, data harmonization, disclosure, and transferring issues with the data required (Source 1).

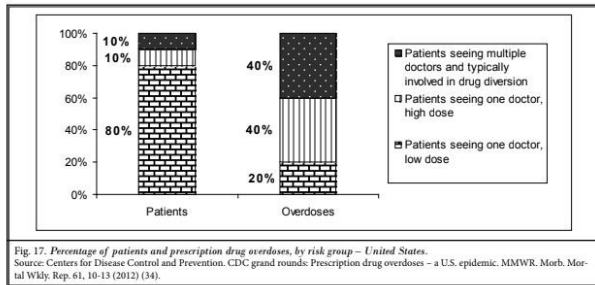


Figure 4: Graphic highlighting percentage of patients and prescription drug overdoses (Manchikanti, 2012).

Discussion and Conclusions

This case study set out to understand the NHCS, its goals, and its purpose. The literature review cited widespread EHR adoption as a reason to leverage the vast data housed therein utilizing tools, such as HL7, to increase secondary use of EHR data. The key informant discussions revealed diverse opinions of how to share EHR data. The informants were in agreement on importance of sharing the data housed in EHRs. The NHCS goals were found to be immediately relevant, but not quickly attainable. Nationally representative data is necessary for proper tracking of phenomena like the opioid crisis. The NHCS has not yet met their recruitment goals in order to collect this data (Source 1).

The interoperability problems between EHRs and other software's are well known. The process of electronic data transmission into surveys relies on the further standardization and easier extraction of data from EHR systems. This may be possible through data standards, such as HL7 standards, specific to NHCS. The affordability of certain solutions, such as HL7, was called into question by both literature and interviews. There are multiple methodologies implemented to work through reporting on and using EHR data, all with their own costs and issues.

The switching of data collection from manual surveys to EHRs intends on reducing stakeholder burden, and increasing time allocated for patient care (Christino, 2013). It should be faster and more secure, as well as provide more information for the purpose of reviewing national health statistics (Brown, 2015). HL7 is expected to improve as physician interaction with EHRs are better understood and data capture will reap the benefits (Dolin, 2006). Secondary use of EHR data is thought to be a huge stepping stone in enhancing healthcare both in the clinic and outside of the clinic. A suggestion made to increase interoperability would be an agreed-upon "essential content for EHR", which would make future secondary use simpler and more accurate (Botis, 2010).

The NHCS has the potential to track the impact of Accountable Care Organizations (ACO) and EHR imple-

mentation on delivery and quality of healthcare and clinical decision support. Uniform Billing (UB)-04 data, which are the billing documents sent to and reviewed by the Center for Medicaid and Medicare Services (CMS), for the purpose of better tracking and comparing billing across hospitals, is collected by the survey (DeFrances, 2012). The UB-04 data allows for reporting important information to participants regarding their readmission rates, 30-day mortality rates, rates of hospital acquired conditions, length of inpatient stays, and intensive care use (DeFrances, 2012). Electronic quality measure survey reporting may increase widespread information on hospital safety without additional time burden on clinic staff (Weiner, 2012).

This study was limited in scope, given that the NHCS was the first national healthcare survey to be made electronic. The novelty of the data collection methods limited the available literature. The HL7 IG did not conduct a roll-out to be measured and it is not yet widely adopted by EHR vendors (Source 1). The NHCS itself has not conducted validation testing, further limiting the ability to fully understand the feasibility for electronic healthcare surveys (Source 1). This study was also not able to include testimony by NHCS HL7 IG EHR vendors, to receive cost and upload details necessary to determine feasibility. Finally, there were not enough key informants interviewed in order to obtain well-rounded expert opinions.

In the future, this review may serve as an example of the discussion surrounding other large public health surveys being replaced by leveraging clinical data from EHRs. The challenges of trying to leverage EHR data offer new opportunities for Informaticists. It is a technically complicated task to collect electronic healthcare data securely and in a standardized manner. It is important for Informaticists to find safe and efficient ways to use EHR data for secondary uses to further public health research and statistics.

Acknowledgements

This work would not have been possible without the creativity and mentorship of Dr. Prashila Dullabh. Continued support and advice from Dr. David Liebovitz have helped make this paper and future work possible. Many of the contacts made and attempted were possible only through the help of Dr. Suzanne Cox and Imran Khan. John Moses was extremely helpful in taking the time to explain the intricacies of EHR reporting and data extraction. Abigail Watson provided expert support on HL7 for this project. Finally, Dr. Vojislav Pejovic went above and beyond to help make this paper of quality.

References

Section: EHR Data

- 1) Wagaw, F., Okoro, C. A., Kim, S., Park., Rachman, F., (2018). Linking Data from Health Surveys an Electronic Health Records: A Demonstration Project in Two Chicago Health Center Clinics. *Preventing Chronic Disease*, 2018, 15:09. Retrieved July 22, 2018.
- 2) Scholte, M., van Dulmen, S. A., Neeleman-Van der Steen, C. W. M., van der Wees, P. J., Nijhuis-van der Sanden, M. W. G., Braspenning, J., (2016). Data extraction from electronic health records (EHRs) for quality measurement of the physical therapy process: comparison between EHR data and survey data. *BMC Medical Informatics and Decision Making*, 2016, 16:141. Retrieved July 25, 2018.
- 3) Murdoch, T. B., Detsky, A. S. (2013), The Inevitable Application of Big Data to Health Care. *Journal of the American Medical Association*, 309(13), 1351-1352. Retrieved February 04, 2018.
- 4) Violan, C., Foguet-Boreu, Q., Hermosilla-Perez, E., Valderas, J. M., Bolibar, B., Fabregas-Escuriola, M., Brugulat-Guiteras, P., Munoz-Perez, M. A., (2013). Comparison of the information provided by electronic health records data and a population health survey to estimate prevalence of selected health conditions and multimorbidity. *BMC Public Health*, 2013, 13:251. Retrieved July 21, 2018.
- 5) Weiner, J. P., Fowles, J. B., Chan, K. S., (2012). New paradigms for measuring clinical performance using electronic health records. *International Journal for Quality in Health Care*, 2012, 24:3, 200-206. Retrieved July 25, 2018.
- 6) Rajeev, D., Staes, C. J., Evans, S., Mottice, S., Rolfs, R., Samore, M. H., Whitney, J., Kurzban, R., Huff, S. M. (2010), Development of an electronic public health case report using HL7 v2.5 to meet public health needs. *Journal American Medical Informatics Association*, 17, 34-41. Retrieved February 04, 2018.
- 7) Botsis, T., Hartvigsen, G., Chen, F., & Weng, C. (2010). Secondary Use of EHR: Data Quality Issues and Informatics Opportunities. *Summit on Translational Bioinformatics*, 2010, 1-5. Retrieved July 1, 2018.
- 8) Chan, K. S., Fowles, J. B., Weiner, J. P. (2009), Electronic Health Records and the Reliability and Validity of Quality Measures: A Review of the Literature. *Medical Care Research and Review SAGE*, 67(5), 503-527. Retrieved February 04, 2018.
- 9) Safran, C., Bloomrosen, M., Hammond, E., Labkoff, S., Markel-Fox, S., Tang, P. C., Detmire, D. E. (2007), Toward a National Framework for the Secondary Use of Health Data: An American Medical Informatics Association White Paper. *Journal of the American Medical Informatics Association*, 14, Jan/Feb. Retrieved February 12, 2018.
- 10) Dolin, R.H., Alshuler, L., Boyer, S.L., Beebe, C., Behlen, F.M., Biron, P.V., Shabo, A., (2006), The HL7

Clinical Document Architecture, Release 2, *JAMA*, 13(1), 30-39. Retrieved June 18, 2018.

- 11) Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., Taylor, R., (2005), Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, and Costs, *EMR Systems*, 24(5), 1103-1117. Retrieved June 24, 2018.

- 12) Dolin, R.H., Alshuler, L., Beebe, C., Biron, P.V., Boyer, S.L., Essin, D., Kimber, E., Lincoln, T., Mattison, J.E., (2001), The HL7 Clinical Document Architecture, *JAMA*, 8(6), 552-569. Retrieved June 18, 2018.

Section: Healthcare Regulations

- 13) Adler-Milstein, J., Bates, D. W., Jha, A. K., (2011). A Survey of Health Information Exchange Organizations in the United States: Implications for Meaningful Use. *Annals of Internal Medicine*, 2011, 154:666-671. Retrieved July 6, 2018.

- 14) Blumenthal, D., Tavenner, M., (2010). The "Meaningful Use" Regulation for Electronic Health Records. *The New England Journal of Medicine*, 363:6, 501-504. Retrieved July 6, 2018.

- 15) The Office of the National Coordinator for Health Information Technology (2017), <https://dashboard.healthit.gov/apps/health-information-technology-data-summaries.php?state=National&cat9=all+data#summary-data>. Retrieved February 24, 2018.

- 16) Health Level Seven International (2017), http://www.hl7.org/implement/standards/product_brief.cfm?product_id=385, Retrieved August 8, 2018.

- 17) CDC, (2018), Updated Declaration of Readiness for the National Health Care Surveys. February 7, 2018. https://www.cdc.gov/nchs/data/nhcs/declaration_of_readiness_nhcs. Retrieved August 27, 2018.

Section: Healthcare Surveys

- 18) DeFrances, C., Lau, D., (2018), Collecting Electronic Health Record Data for the National Ambulatory Medical Care Survey (NAMCS) and the National Hospital Care Survey (NHCS). *Presentation to the Federal Committee on Statistical Methodology 2018 Research and Policy Conference*, March 9, 2018. Retrieved August 8, 2018.

- 19) DeFrances, C., (2017), Update on the EHR Data Collections for the National Ambulatory Medical Care Survey (NAMCS) and the National Hospital Care Survey (NHCS). *Presentation to the NCHS Board of Scientific Counselors September 6, 2017*. Retrieved August 24, 2018.

- 20) Brown, D., (2015), Use of Electronic Health Record Data Collection by the National Health Care Surveys. *National Academy of Sciences Workshop on Use of Alternate and Multiple Data Sources for Federal Studies*, December 16, 2015. Retrieved August 8, 2018.

- 21) Christino, M. A., Matson, A. P., Fischer, S. A., Reinert, S. E., DiGiovanni, C. W., Fadale, P. D., (2013), Paperwork

Versus Patient Care: A Nationwide Survey of Residents' Perceptions of Clinical Documentation Requirements and Patient Care, *Journal of Graduate Medical Education*, December 2013, 600-604. Retrieved February 24, 2018.

22) DeFrances, C., Levant, S., (2012), Electronic Collection of Inpatient and Ambulatory Hospital Care Data: National Hospital Care Survey. *The Proceedings of the 13th Annual International Conference on Digital Government Research*, June 4-7 2012. Retrieved February 04, 2018.

Section: Opioid Epidemic

23) Brown, A. M., DeFrances, C., Crane, E., Cai, R., Naeger, S., (2018), Identification of Substance-involved Emergency Department Visits Using Data From the National Hospital Care Survey. *National Health Statistics Reports*, August 14, 2018, Number 114. Retrieved October 10, 2018.

24) Manchikanti, L., Helm II, S., Fellows, B., Janata, J. W., Pampati, V., Grider, J. S., Boswell, M. V., (2012), Opioid

Epidemic in the United States. *Pain Physician Journal*, July 12, 15:ES9-ES38. Retrieved September 27, 2018.

Section: Key Informants

25) Source 1: Dr. Carol DeFrances Branch Chief, Ambulatory and Hospital Care Statistics Branch of the US Department of Health and Human Services. September 25, 2018. Phone call. 12.

26) Source 2: Dr. David Liebovitz the Chief Medical Information Office of the University of Chicago Medicine. September 21, 2018. In-person. 13.

27) Source 3: John Moses, Director of Enterprise Architecture and New Technologies for University of Chicago Medicine. October 5, 2018. Webex. 14.

28) Source 4: Abigail Watson, Biomedical Informatics Graduate University of Chicago. September 28, 2018 - November 02, 2018. Emails.