

2-23-2017

Active Use of Electronic Health Records (EHRs) and Personal Health Records (PHRs) for Epidemiologic Research: Sample Representativeness and Nonresponse Bias in a Study of Women During Pregnancy

Julie K. Bower

Ohio State University - Main Campus, bower.185@osu.edu

Claire E. Bollinger

The Ohio State University, bollinger.69@osu.edu

Randi E. Foraker

The Ohio State University, rforaker@cph.osu.edu

Darryl B. Hood

The Ohio State University, hood.188@osu.edu

See next pages for additional authors

Follow this and additional works at: <http://repository.edm-forum.org/egems>

Recommended Citation

Bower, Julie K.; Bollinger, Claire E.; Foraker, Randi E.; Hood, Darryl B.; Shoben, Abigail B.; and Lai, Albert M. (2017) "Active Use of Electronic Health Records (EHRs) and Personal Health Records (PHRs) for Epidemiologic Research: Sample Representativeness and Nonresponse Bias in a Study of Women During Pregnancy," *eGEMs (Generating Evidence & Methods to improve patient outcomes)*: Vol. 5: Iss. 1, Article 1.

DOI: <http://dx.doi.org/10.13063/2327-9214.1263>

Available at: <http://repository.edm-forum.org/egems/vol5/iss1/1>

This Methods Case Study is brought to you for free and open access by the the Publish at EDM Forum Community. It has been peer-reviewed and accepted for publication in eGEMs (Generating Evidence & Methods to improve patient outcomes).

The Electronic Data Methods (EDM) Forum is supported by the Agency for Healthcare Research and Quality (AHRQ), Grant 1U18HS022789-01. eGEMs publications do not reflect the official views of AHRQ or the United States Department of Health and Human Services.

Active Use of Electronic Health Records (EHRs) and Personal Health Records (PHRs) for Epidemiologic Research: Sample Representativeness and Nonresponse Bias in a Study of Women During Pregnancy

Abstract

Introduction: With the growing use of electronic medical records, electronic health records (EHRs), and personal health records (PHRs) for health care delivery, new opportunities have arisen for population health researchers. Our objective was to characterize PHR users and examine sample representativeness and nonresponse bias in a study of pregnant women recruited via the PHR.

Design: Demographic characteristics were examined for PHR users and nonusers. Enrolled study participants (responders, n=187) were then compared with nonresponders and a representative sample of the target population.

Results: PHR patient portal users (34 percent of eligible persons) were older and more likely to be White, have private health insurance, and develop gestational diabetes than nonusers. Of eligible persons (all PHR users), 11 percent (187/1,713) completed a self-administered PHR-based questionnaire. Participants in the research study were more likely to be non-Hispanic White (90 percent versus 79 percent) and married (85 percent versus 77 percent), and were less likely to be Non-Hispanic Black (3 percent versus 12 percent) or Hispanic (3 percent versus 6 percent). Responders and nonresponders were similar regarding age distribution, employment status, and health insurance status. Demographic characteristics were similar between responders and nonresponders.

Discussion: Demographic characteristics of the study population differed from the general population, consistent with patterns seen in traditional population-based studies. The PHR may be an efficient method for recruiting and conducting observational research with additional benefits of efficiency and cost-effectiveness.

Acknowledgements

This study was funded by an a seed grant from the Ohio State University College of Public Health.

Keywords

Electronic Health Records, Data Collection, Population Health, Epidemiology, Bias

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](https://creativecommons.org/licenses/by-nc-nd/3.0/).

Authors

Julie K Bower, *Ohio State University - Main Campus*; Claire E Bollinger, *The Ohio State University*; Randi E Foraker, *The Ohio State University*; Darryl B Hood, *The Ohio State University*; Abigail B Shoben, *The Ohio State University*; Albert M Lai, *Ohio State University - Main Campus*.



Active Use of Electronic Health Records (EHRs) and Personal Health Records (PHRs) for Epidemiologic Research: Sample Representativeness and Nonresponse Bias in a Study of Women During Pregnancy

Julie K Bower, PhD, MPH;ⁱ Claire E. Bollinger;ⁱⁱ Randi E. Foraker, PhD;^j Darryl B. Hood, PhD;ⁱⁱ Abigail B. Shoben, PhD;ⁱⁱⁱ Albert M. Lai, PhD^{iv}

ABSTRACT

Introduction: With the growing use of electronic medical records, electronic health records (EHRs), and personal health records (PHRs) for health care delivery, new opportunities have arisen for population health researchers. Our objective was to characterize PHR users and examine sample representativeness and nonresponse bias in a study of pregnant women recruited via the PHR.

Design: Demographic characteristics were examined for PHR users and nonusers. Enrolled study participants (responders, n=187) were then compared with nonresponders and a representative sample of the target population.

Results: PHR patient portal users (34 percent of eligible persons) were older and more likely to be White, have private health insurance, and develop gestational diabetes than nonusers. Of eligible persons (all PHR users), 11 percent (187/1,713) completed a self-administered PHR based questionnaire. Participants in the research study were more likely to be non-Hispanic White (90 percent versus 79 percent) and married (85 percent versus 77 percent), and were less likely to be Non-Hispanic Black (3 percent versus 12 percent) or Hispanic (3 percent versus 6 percent). Responders and nonresponders were similar regarding age distribution, employment status, and health insurance status. Demographic characteristics were similar between responders and nonresponders.

Discussion: Demographic characteristics of the study population differed from the general population, consistent with patterns seen in traditional population-based studies. The PHR may be an efficient method for recruiting and conducting observational research with additional benefits of efficiency and cost-effectiveness.

ⁱDivision of Epidemiology, The Ohio State University College of Public Health, ⁱⁱDivision of Environmental Health Sciences, The Ohio University College of Public Health, ⁱⁱⁱDivision of Biostatistics, The Ohio State University College of Public Health, ^{iv}Institute for Informatics, Washington University School of Medicine

Introduction

The electronic health record (EHR) contains extensive information collected by clinicians about a patient's health status, compiled information from various health care providers, and encounters with the health care system for a particular patient. The EHR can be queried to quickly identify patients that meet specific inclusion and exclusion criteria for epidemiologic research purposes. Personal Health Records (PHRs) are patient-facing platforms that allow patients to interface with their EHR. With the growing use of EHRs and PHRs for health care delivery and quality improvement, new opportunities also have arisen for rapidly identifying, recruiting, and collecting contextual data from patients for population health research.¹

PHRs have been evaluated in the literature from patient and provider perspectives,² but little is known regarding their utility for conducting epidemiologic research. From a provider perspective, PHRs can assist in population care management between clinical encounters.^{3,4} Of particular concern to epidemiologic researchers is the threat of selection bias—issues of sample representative and nonresponse—given that PHRs may be utilized inconsistently across a patient population and that conclusions drawn based on data collected from PHR users may not be generalizable to a specified target population.⁵ Nonetheless, PHR use is increasingly ubiquitous; patient characteristics currently associated with greater PHR use include the following: younger age, White race, female sex, and greater health care utilization (e.g., patients with multiple and chronic conditions, patients with more comprehensive preventive care coverage).⁶⁻⁸ Based on frequent health care system encounters commonly associated with prenatal care, pregnant women may also fit into this latter category.

Use of the EHR for research may be termed active or passive. Passive use includes activities such as secondary data analysis or identification of patients that meet specific phenotypic requirements. Active use includes engaging patients via the use of platforms such as the EHR and PHR for recruitment and outcomes data collection.^{9,10} The overall goal of this study is to describe the use, strengths, and limitations of the EHR for identifying potential study participants and the PHR for recruiting and collecting contextual data for an observational research study. We describe demographic characteristics of a population of PHR users and nonusers in a large academic health center. We then compare demographics of a study population recruited through the PHR to two comparison populations: (1) nonresponders (i.e., individuals invited to participate but that did not complete data collection activities), and (2) a representative probability sample of pregnant women in a specified target population.

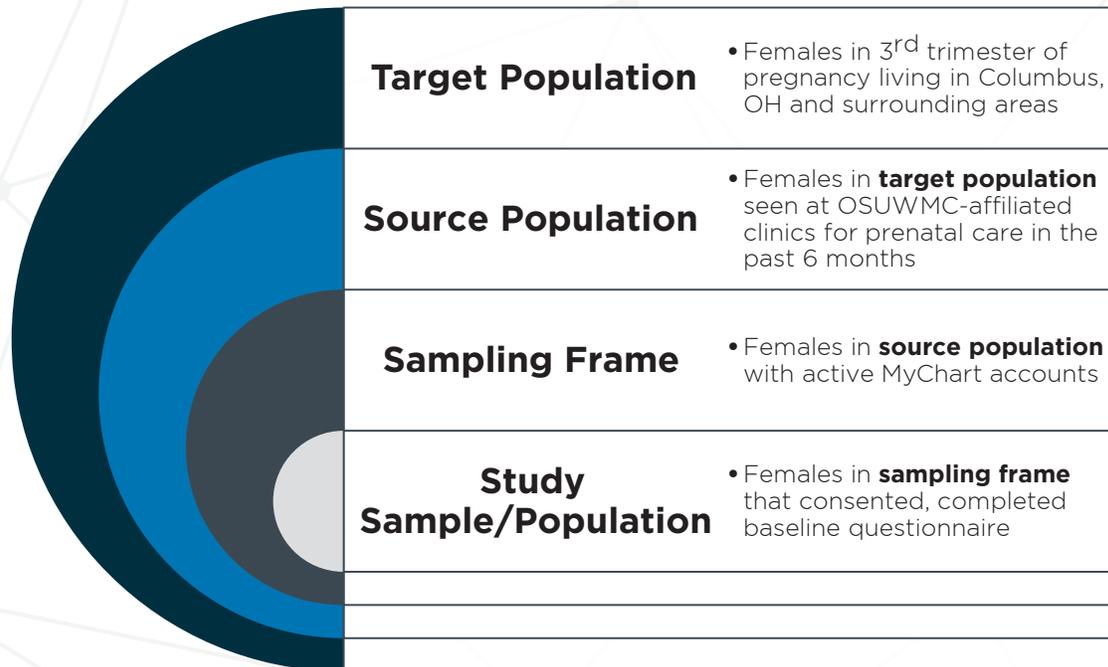
Design

Comparison of PHR Users Versus Nonusers: Characterizing the Sampling Frame

In preparation for study implementation, we queried all pregnancies that met the eligibility criteria during a one-year period (November 1, 2013 to October 31, 2014) to characterize the representativeness of PHR “users” (the sampling frame, defined as the list of patients who would be eligible to participate and have access to the PHR-based questionnaire) to the target population of pregnant women in the catchment area (Figure 1). The following data were extracted from the EHR: PHR account activation status (none, patient declined, activated but not used, activated), age, race, ZIP code of primary residence, primary health insurance type, history of diabetes, gestational diabetes diagnosis any time during most recent pregnancy, and preeclampsia diagnosis any time during the current pregnancy.



Figure 1. Summary of Sampling Approach



Study Population

To illustrate the utility of the EHR and PHR for observational research and data collection, we use data from a study that was designed to assess associations between clinical factors, environmental exposures, health behaviors, and community-built environment features with metabolic and cardiovascular outcomes in a population of pregnant women. One hundred eighty-seven women being seen within The Ohio State University Wexner Medical Center (OSUWMC) affiliate system in the Columbus, Ohio metropolitan area were recruited during their third trimester of pregnancy over a six-month period (March 31–September 30, 2015) into an observational research study. Participants completed a self-reported questionnaire and were followed for the duration of their current pregnancy. The Ohio State University Biomedical Institutional Review Board reviewed and approved the research protocol prior to the start of data collection.

Implementation

The OSUWMC uses the Epic (Verona, WI) EHR system. Patients were recruited via Epic's patient-facing PHR portal, MyChart. The EHR at OSUWMC was queried monthly to identify women seen in the prior six months with the following inclusion criteria: currently pregnant, third trimester (≥ 28 weeks), active MyChart account (defined as "opting into the platform, registering, and logging in at least once"), and not deceased. During the data collection period, 1,713 eligible patients were identified, and all received the same recruitment message via their MyChart inbox. Additionally, a message was sent to the email address with which patients registered for their account alerting them that their MyChart account contained new content. Information about the study and research team was provided; interested patients then electronically consented to participate in the research study. Prior to the questionnaire start, eligibility criteria were verified via self-report.

Participants (n=187) completed a 110-item self-administered questionnaire. Participants were able to decline further participation and to delete responses at any time in the event that they no longer wanted to contribute data. All data were collected via MyChart and stored behind the OSUWMC firewall; completed questionnaire responses were stored in a secured data mart for later retrieval by the study team. Following the participant's delivery date, demographic and clinical information were abstracted and linked with questionnaire responses. "Responders" were defined as those women that completed the MyChart questionnaire. Note that, in the version of MyChart used, participants were not able to use the mobile app to complete the questionnaire and were required to use the desktop web version of MyChart, as the MyChart mobile app did not yet support questionnaire functionality at the time of this study.

Comparison of Responders Versus Nonresponders: Characterizing the Study Sample

During the active data collection period, the following de-identified data were stored for all potential patients that were invited to participate: last MyChart login date, age, race, ethnicity, ZIP code, and estimated delivery date. These patients were defined as the "sampling frame" for the study; those that completed the questionnaire were the "responders" and study population. We compared key demographic characteristics of responders (i.e., study participants) with the sampling frame (i.e., all those that met eligibility criteria and had an active MyChart account). Additionally, in order to describe the representativeness of the study population to the target population of pregnant women in the Columbus, Ohio area, we compared key demographics to aggregate publicly available demographic data previously reported by the Pregnancy Risk Assessment Monitoring System (PRAMS) surveillance project, a population-

based survey conducted by the Ohio Department of Health.^{11,12} PRAMS data are self-reported by approximately 200 women per month, and are compiled to supplement data from birth certificates and to produce generalizable estimates on all live births in the state of Ohio.¹² We selected Ohio Region 4 to match the catchment area for this project.

EHR Data

For responders who agreed to participate in the study and contribute self-reported data and clinical information, the following data were collected from the EHR following the patient's delivery date: demographics (e.g., age, race, ethnicity, smoking history, health history); health status (e.g., new diagnoses during pregnancy); laboratory and clinical tests (e.g., blood glucose measurements, blood pressure, height, weight); and delivery information (e.g., birth outcome, complications during delivery, discharge diagnoses). For patients that did not complete the MyChart questionnaire (nonresponders), only demographics were collected from the EHR and are included in the current study.

Statistical Analysis

We compared demographics between MyChart "users" (n=1,977 over a one-year period) and "nonusers" (n = 3,782 over a one-year period) to characterize the representativeness of the MyChart users to the target population of all pregnant women in the catchment area actively receiving prenatal care. We then compared patient demographics between questionnaire responders (n=187) and nonresponders (i.e., those patients that were invited to participate but declined or did not view the recruitment message, n=1,528) to describe representativeness of the study population to the sampling frame of all eligible patients who were MyChart users; Chi square tests and logistic regression were used to examine factors associated with likelihood of participation. For purposes of



describing the potential representativeness of our study sample, we then compared respondents' demographics to aggregate data for the Columbus, Ohio area collected by the PRAMS.¹³ Analyses were performed using Stata Statistical Software: Release 13.1 (College Station, Texas: StataCorp LP).

Findings

The sampling approach for the current study is summarized in Figure 1. Our target population included all women in the third trimester of pregnancy living in the Franklin County, Ohio area. The source population included the subset of all pregnant women in the target population that were seen at OSUWMC-affiliated clinics for prenatal

care. Within this population, all women with active MyChart accounts (the sampling frame) were invited to participate in the study. Finally, 187 women were included in the study sample and represented those women in the sampling frame that consented to participate in the PHR-based study and completed the baseline questionnaire.

Comparison of PHR Users Versus Nonusers: Characterizing the Sampling Frame

Approximately 6,000 babies are delivered at OSUWMC each year. Table 1 summarizes key demographic characteristics of mothers of all live births over a one-year period. Of the 5,759 live births from the period October 2013–November

Table 1. Comparison of MyChart Users Versus Nonusers (n=5,759)

	TOTAL <i>n</i> = 5,789	MyChart USERS <i>n</i> = 1,977	MyChart NONUSERS <i>n</i> = 3,782	<i>P</i>
Age, %				< 0.001
< 20 years	4.5	0.8	6.5	
20–24 years	17.6	8.4	22.4	
25–34 years	60.3	67.6	56.5	
≥ 35 years	17.6	23.3	14.6	
Race, %				<0.001
White	61.0	74.1	54.2	
Black or African American	22.5	10.8	28.6	
Asian	3.1	5.4	1.9	
Other	12.6	9.1	14.4	
More than one race	0.8	0.8	0.9	
Health insurance, %				<0.001
Private	84.3	96.8	77.8	
Public	15.7	3.2	22.2	
Other/Uninsured	0.1	0	0.1	
History of diabetes, %	3.0	2.6	3.2	0.2
Gestational diabetes, %	6.5	7.9	5.7	0.002

2014, 34 percent of patients were MyChart and PHR users. Those more likely to use MyChart were women who were aged ≥ 35 years, White, and who had private health insurance. The prevalence of gestational diabetes was higher among MyChart users compared to nonusers (7.89 percent versus 5.74 percent, $P < 0.01$), but the groups were similar regarding history of diabetes prior to the pregnancy.

Comparison of Responders Versus Nonresponders: Characterizing the Study Sample

During the data collection period, we queried the EHR monthly to identify patients newly meeting the eligibility criteria. Of the 1,713 eligible patients identified and invited to participate, 187 participants (11 percent) enrolled in the study. Responders (compared to nonresponders) were more likely to be non-Hispanic White in race or ethnicity (90 percent versus 79 percent) and married (85 percent versus 77 percent). Responders were less likely to be non-Hispanic black (3 percent versus 12 percent) or Hispanic (3 percent versus 6 percent, Table 2) than were nonresponders. In logistic regression models, none of these factors were independently associated with the likelihood of responding to the questionnaire (data not shown). There were no significant differences between the two groups with respect to age, employment status, or health insurance type.

Comparison of Responders Versus Ohio PRAMS Region 4: Representativeness of the Study Sample

Compared to population-based data on pregnancies in the catchment area from the 2009–2011 PRAMS, responders had a higher representation of non-Hispanic White women, higher proportion of women with educational attainment beyond a high school diploma, and comprised a higher proportion of women who were married (Table 3). Additionally, responders reported a lower prevalence of smoking

in the three months prior to the pregnancy compared to the general population (8 percent versus 34 percent).

Discussion

Prior studies demonstrate the utility and limitations of passive use of administrative health records for secondary data analysis and recruitment of population-based samples.^{14–18} Our study extends this work to evaluate the use of the PHR for data collection activities. Overall, we found that our study sample was one that did not mirror the general population regarding distribution of demographic factors but these differences were similar to those observed in traditional observational epidemiologic studies. However, because we had extensive information about nonresponders and non-PHR users, differences could be more completely described, quantified, and directly compared to various external populations toward more accurate interpretation of the generalizability and representativeness of observed study data. One limitation of our study to note is that we collected data about PHR users from only one health system. Thus, we were unable to formally evaluate how the patient population described in this study differs from other health system patient populations.

One of the chief strengths of using the PHR for data collection over traditional methods such as mailed surveys is access to extensive data through the EHR to characterize nonresponders and to quantify the magnitude and direction of potential nonresponse bias. In the current study, we present an example comparing key demographic characteristics of responders versus nonresponders to a self-administered questionnaire, MyChart users versus nonusers, and our study sample of responders to the general and target population. These rich data sources allow us to draw stronger inferences about the target population based on our study sample,



Table 2. Comparison of Responders to Nonresponders (n = 1,713)

	TOTAL <i>n</i> = 1,713	RESPONDERS <i>n</i> = 187	NONRESPONDERS^a <i>n</i> = 1,528	<i>P</i>
Age, %				0.2
< 20 years	0.8	0	0.9	
20–24 years	9.0	5.4	9.4	
25–34 years	65.9	68.5	65.6	
≥ 35 years	24.3	26.2	24.1	
Race, %				0.001
Non-Hispanic White	80.4	89.7	79.2	
Non-Hispanic Black	10.7	3.3	11.7	
Hispanic	5.8	2.7	6.2	
Non-Hispanic Other	3.1	4.4	3.0	
Employment status, %				0.1
Employed	76.9	82.9	76.1	
Student	2.3	1.6	2.4	
Unemployed/unknown	20.9	15.5	21.5	
Marital status, %				
Married	78.8	85.0	77.4	0.02
Not married or other	21.8	15.1	22.6	
Insurance status ^b , %				0.8
Private	96.5	96.5	96.8	
Public/government/other	3.5	3.2	3.5	

Notes: ^a“Nonresponders” were defined as “individuals with active MyChart accounts that received a recruitment message but did not complete the questionnaire in response to a single recruitment message.”

^bStatus at most recent visit during the eligibility window.

Table 3. Comparison of OSUWMC Study Sample Compared to Ohio PRAMS¹³

	OSUWMC (APRIL–AUGUST 2015, n=187)		COLUMBUS REGION 4 (2009–2011) ^a	
	PERCENT	95% CI	PERCENT	95% CI
Age, years (n=114)				
Less than 20	0		9.8	(7.7, 12.3)
20–24	5.4	(2.1, 8.6)	26.6	(23.3, 30.1)
25–34	68.5	(61.7, 75.5)	51.3	(7.6, 55.0)
35 or older	26.2	(19.8, 32.6)	12.4	(10.2, 14.9)
Race				
Non-Hispanic White	89.7	(85.2, 94.1)	76.4	(73.5, 79.1)
Non-Hispanic Black	3.3	(0.7, 5.9)	13.3	(11.7, 15.1)
Hispanic	4.3	(1.4, 7.3)	2.8	(1.7, 4.6)
Non-Hispanic Other	2.7	(0.4, 5.1)	7.6	(5.8, 9.8)
Education, years				
Less than 12	0		19.7	(16.7, 23.2)
12	1.8	(0, 3.8)	23.3	(20.2, 26.6)
More than 12	98.2	(96.2, 100)	57.0	(53.2, 60.8)
Marital Status				
Married	84.5	(79.0, 90.0)	54.3	(50.6, 58.0)
Other	15.5	(10.0, 21.0)	45.7	(42.0, 49.5)
Mothers smoked in the three months before pregnancy ^b	8.3	(4.1, 12.5)	34.0	(30.4, 37.7)

Notes: ^aWeighted population estimates.^bSelf-reported.

and to potentially statistically adjust our estimates to match various target populations through the use of methods such as weighting of observations in our analytic data set. Conducting data collection activities within the PHR allows us to address limitations of passive use of administrative records, namely in collecting additional information that may be missing or difficult to locate in the EHR alone.

A second benefit of leveraging the PHR for data collection activities is efficiency. Once inclusion and exclusion criteria are established, identification of eligible potential participants can be automated and a prescreening query can be written to do so in real time; this is particularly beneficial for identifying rare phenotypes or capturing information during key time windows (in this case, the third trimester



of pregnancy). Additionally, this method is ideal for time-sensitive issues where rapid identification and recruitment of patients meeting specific inclusion and exclusion criteria is critical.

Once eligibility is determined, automated recruitment messages can be sent directly to patients via their PHR—where they receive other communications from their health care providers. Using this secure platform to recruit patients with particular phenotypes (i.e., meeting specific inclusion and exclusion criteria based on individual and clinical characteristics) is scalable, efficient, and protects the privacy of patients' health information. Additionally, patients can obtain information from their PHR to learn more about the study details and investigators at their own convenience, and within this environment can directly contact their health care provider to solicit guidance to inform their decision about whether or not to participate in the study.

Finally, patient portals can be leveraged for collection of patient-reported outcomes (PRO) data that are not otherwise available in the EHR, improving the applicability of using the PHR for public health research. The EHR is designed to aid patient management and clinical decision-making but lacks detailed contextual information such as information about a patient's health behaviors and social context.^{19,20} These variables are essential for studies in which population health researchers examine clinical variables in the context of patient characteristics, attitudes, perceptions, and behaviors. Patient portals such as MyChart allow for efficient collection of important PRO data that can then be combined with phenotype and clinical information found in the EHR as well as data on health care utilization patterns (e.g., use of outpatient and inpatient services, cost of care).

One factor that may have limited our observed response rate may be the device used for MyChart

access. Non-Hispanic black and Hispanic adults are more likely to own a smartphone and are more likely to be smartphone dependent for broadband Internet access than are White adults.²¹ In the version of MyChart used, participants were not able to use the mobile app to complete the questionnaire and were required to use the desktop web version of MyChart, as the MyChart mobile app did not yet support questionnaire functionality. As this capability improves and uptake of PHRs increases, a reduction in selection bias regarding study participants recruited via this method should occur.

Recommendations

Data collected from a sample of a target population can lead to biased study results because we rarely obtain a simple random sample of our target population when active participation is required from individuals (e.g., completion of a questionnaire or a physical examination).²² Estimating representativeness of a study sample has been accomplished in prior studies using various approaches. The first is to assume a simple random sample of the target population. This assumption is rarely met, but obtaining a nonrepresentative sample is not always a threat to study validity or generalizability.²³ For example, the Framingham Heart Study was one of the first to note a (true) positive association between smoking and risk of heart disease.²⁴ This finding was observed in a population of predominately non-Hispanic White men and women enrolled between the ages of 30 and 62 years, all living in the town of Framingham, Massachusetts. This finding has been consistently replicated in other populations since the biological basis of the association between tobacco and heart disease is similar across population subgroups. However, when estimating the burden of health conditions such as the prevalence of cardiovascular disease in the general population, obtaining a representative sample is of key importance.

Rather than assuming no bias, we can try to estimate the bias introduced by not having the participation of nonresponders. One way to do this is to compare key measures in the study population (e.g., demographics) with those of an appropriate comparison population. However, identifying another study with an appropriate comparison population is not always possible. Thus, a second approach is to use information from prior literature to make informed assumptions about characteristics of the missing respondents or to compare available demographics (e.g., age, sex, and race or ethnicity distribution) between responders and nonresponders. This is ideal since the comparison data come from the population of interest, but these data are not always available. When they are, we might have access only to limited demographic information—and assumptions are subsequently made that the associations found in the study population would be the same as if nonrespondents had participated—given that they are similar to those with matching demographic characteristics.

Using the EHR allows for more accurate estimation of nonresponse bias because a comprehensive list of information about nonresponders can be directly obtained from the EHR, including health status information. When sociodemographic information is relevant, data on factors such as health insurance status and health care utilization patterns can be categorized.

Prior literature suggests that delayed responders—those that agree to participate later rather than earlier for a given research study in response to a reminder postcard or similar prompt—will be similar to nonrespondents.^{25,26} Thus, a third approach to estimating representativeness is to compare key participant characteristics between early responders and delayed responders. If obtaining a representative sample is important to answer a particular research question, employing similar strategies using the

patient portal or following up via phone or mail is recommended. Additionally, incentives were not used in our study but might be considered to improve response rates in future studies. Finally, weighting approaches can be employed to make more accurate estimates from the observed study data using a vast amount of known data passively extracted from the EHR.

Conclusion

The EHR offers advantages to patient care, consolidates information on a patient's experience with the health care system in a centralized location, and improves coordination and continuity of care. Passive EHR use for research purposes is accelerating; active use of the PHR is still emerging and has understandably been approached with caution. This study demonstrates that the EHR and PHR can be leveraged for research study eligibility screening, recruitment, and data collection to enhance research that relies upon the collection of additional information not typically collected during a patient encounter in the inpatient or outpatient setting.

Merging of PRO data obtained through the PHR with a patient's EHR record allows researchers to address questions that otherwise could not be answered with the EHR alone. For epidemiologic researchers, the EHR further provides a rich data source for characterizing individuals that are typically missed in research studies because of nonparticipation, ultimately improving the ability to use observational data from a study sample to answer important scientific questions about the larger population. As the digital divide continues to narrow, the inclusion of underrepresented population subgroups should also increase to further strengthen inferences from population-based studies using this existing administrative infrastructure.



Acknowledgements

This study was funded by a seed grant from the Ohio State University College of Public Health.

References

1. Kite BJ, Tangasi W, Kelley M, Bower JK, Foraker RE. Electronic Medical Records and Their Use in Health Promotion and Population Research of Cardiovascular Disease. *Current Cardiovascular Risk Reports*. 2015;9(1):1-8.
2. Irizarry T, DeVito Dabbs A, Curran CR. Patient Portals and Patient Engagement: A State of the Science Review. *J Med Internet Res*. 2015;17(6):e148.
3. Price M, Bellwood P, Kitson N, Davies I, Weber J, Lau F. Conditions potentially sensitive to a personal health record (PHR) intervention, a systematic review. *BMC medical informatics and decision making*. 2015;15:32.
4. Gee PM, Paterniti DA, Ward D, Soederberg Miller LM. e-Patients Perceptions of Using Personal Health Records for Self-management Support of Chronic Illness. *Comput Inform Nurs*. 2015 Jun;33(6):229-237.
5. Czaja SJ, Zarcadoolas C, Vaughn WL, Lee CC, Rockoff ML, Levy J. The usability of electronic personal health record systems for an underserved adult population. *Hum Factors*. 2015 May;57(3):491-506.
6. Curtis J, Cheng S, Rose K, Tsai O. Promoting adoption, usability, and research for personal health records in Canada: the MyChart experience. *Healthcare management forum / Canadian College of Health Service Executives = Forum gestion des soins de sante / College canadien des directeurs de services de sante*. 2011 Autumn;24(3):149-154.
7. Gerber DE, Laccetti AL, Chen B, et al. Predictors and intensity of online access to electronic medical records among patients with cancer. *J Oncol Pract*. 2014 Sep;10(5):e307-312.
8. Miller H, Vandenbosch B, Ivanov D, Black P. Determinants of personal health record use: a large population study at Cleveland Clinic. *J Healthc Inf Manag*. 2007 Summer;21(3):44-48.
9. Roumia M, Steinhubl S. Improving cardiovascular outcomes using electronic health records. *Curr Cardiol Rep*. 2014 Feb;16(2):451.
10. Foraker RE, Kite B, Kelley MM, et al. EHR-based Visualization Tool: Adoption Rates, Satisfaction, and Patient Outcomes. *EGEMS (Wash DC)*. 2015;3(2):1159.
11. Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). PRAMS Questionnaires. 2015 July 29, 2015 [cited 2015 August 14]; Available from: <http://www.cdc.gov/prams/questionnaire.htm>
12. Ohio Department of Health. Ohio Pregnancy Risk Assessment Monitoring Program. 2014 11/4/2014 [cited 2015 November 4]; Available from: <https://www.odh.ohio.gov/healthstats/pramshs/prams1.aspx>
13. Pregnancy Risk Assessment Monitoring System, Ohio Department of Health. Ohio PRAMS Perinatal Region Data Summary, 2009-2011; 2014.
14. Oakes JM, MacLehose RF, McDonald K, Harlow BL. Using administrative health care system records to recruit a community-based sample for population research. *Annals of epidemiology*. 2015 Jul;25(7):526-531.
15. Catalan-Ramos A, Verdu JM, Grau M, et al. Population prevalence and control of cardiovascular risk factors: what electronic medical records tell us. *Aten Primaria*. 2014 Jan;46(1):15-24.
16. Danford CP, Navar-Boggan AM, Stafford J, McCarver C, Peterson ED, Wang TY. The feasibility and accuracy of evaluating lipid management performance metrics using an electronic health record. *Am Heart J*. 2013 Oct;166(4):701-708.
17. Hammermeister K, Bronsert M, Henderson WG, et al. Risk-adjusted comparison of blood pressure and low-density lipoprotein (LDL) noncontrol in primary care offices. *Journal of the American Board of Family Medicine : JABFM*. 2013 Nov-Dec;26(6):658-668.
18. Hissett J, Folks B, Coombs L, Leblanc W, Pace WD. Effects of changing guidelines on prescribing aspirin for primary prevention of cardiovascular events. *Journal of the American Board of Family Medicine : JABFM*. 2014 Jan-Feb;27(1):78-86.
19. Weber GM, Mandl KD, Kohane IS. Finding the Missing Link for Big Biomedical Data. *JAMA*. 2014 May 22;310(1001):4228.
20. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA*. 2013 Apr 3;309(13):1351-1352.
21. Pew Research Center. U.S. Smartphone Use in 2015; 2015 April 1, 2015.
22. Szklo M, Nieto FJ. *Epidemiology Beyond the Basics*. 2 ed. Sandbury, MA: Jones and Bartlett Publishers; 2007.
23. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *International journal of epidemiology*. 2013 Aug;42(4):1012-1014.
24. Doyle JT, Dawber TR, Kannel WB, Heslin AS, Kahn HA. Cigarette smoking and coronary heart disease. Combined experience of the Albany and Framingham studies. *The New England journal of medicine*. 1962 Apr 19;266:796-801.
25. Lindner JR, Murphy TH, Briers GE. Handling nonresponse in social science research. *Journal of Agricultural Education*. 2001;42(4):43-53 %@ 1042-0541.
26. Miller LE, Smith KL. Handling nonresponse issues. *Journal of extension*. 1983;21(5):45-50.