INFORMATICS IN PRIMARY CARE

Research article

Can we predict a national profile of non-attendance paediatric urology patients: a multi-institutional electronic health record study

Cite this article: Bush RA, Vemulakonda VM, Corbett ST, Chiang GJ. Can we predict a national profile of non-attendance paediatric urology patients: a multi-institutional electronic health record study. *Inform Prim Care*. 2014;21(3):132–138.

http://dx.doi.org/10.14236/jhi.v21i3.59

Author address for correspondence:
Ruth A. Bush
University of San Diego, San Diego, 92110, USA, Clinical Research Informatics, Rady Children's Hospital-San Diego, 7910 Frost Street Suite #325, San Diego, CA 92123, USA

Accepted April 2014

Ruth A. Bush
University of San Diego, San Diego, 92110, USA
Clinical Research Informatics, Rady Children's Hospital-San Diego, 7910 Frost Street, Suite #325, San Diego, CA 92123, USA

Vijaya M. Vemulakonda
Department of Pediatric Urology, Children's Hospital Colorado, 13123 East 16th Avenue, Box 463, Aurora, CO 80045, USA

Sean T. Corbett
Division of Pediatric Urology, University of Virginia, PO Box 800422, Charlottesville, VA 22908, USA

George J. Chiang
University of California, San Diego, La Jolla, CA 92093, USA
Clinical Research Informatics, Rady Children's Hospital-San Diego, 7910 Frost Street, Suite #325, San Diego, CA 92123, USA

ABSTRACT

Background Non-attendance at paediatric urology outpatient appointments results in the patient's failure to receive medical care and wastes health care resources.

Objective To determine the utility of using routinely collected electronic health record (EHR) data for multi-centre analysis of variables predictive of patient no-shows (NS) to identify areas for future intervention.

Methods Data were obtained from Children's Hospital Colorado, Rady Children's Hospital San Diego and University of Virginia Hospital paediatric urology practices, which use the Epic® EHR system. Data were extracted for all urology outpatient appointments scheduled from 1 October 2010 to 30 September 2011 using automated electronic data extraction techniques. Data included appointment type; date; provider type and days from scheduling to appointment. All data were de-identified prior to analysis. Predictor variables identified using χ^2 and analysis of variance were modelled using multivariate logistic regression.

Results A total of 2994 NS patients were identified within a population of 28,715, with a mean NS rate of 10.4%. Multivariate logistic regression determined that an appointment with mid-level provider (odds ratio (OR) 1.70 95% CI (1.56, 1.85)) and an increased number of days between scheduling and appointment (15–28 days OR 1.24 (1.09, 1.41); 29+ days OR 1.70 (1.53, 1.89)) were significantly associated with NS appointments.

Conclusion We demonstrated sufficient interoperability among institutions to obtain data rapidly and efficiently for use in 1) interventions; 2) further study and 3) more complex analysis. Demographic and potentially modifiable clinic characteristics were associated with NS to the outpatient clinic. The analysis also demonstrated that available data are dependent on the clinical data collection systems and practices.

Keywords: appointments and schedules, electronic health record (EHR), missed appointments, paediatric urology, retrospective study.

What this paper adds:

- Data from the EHR can be used to efficiently identify populations for intervention
- Data can be combined from various institutions to conduct analysis
- Limitations to the robustness of the data can be both technological and regulatory.

INTRODUCTION AND OBJECTIVE

The electronic health record (EHR) was designed as a clinical and administrative tool, and is transforming clinical patient care.[1] The expansive range of information and the sheer volume permitting the identification and evaluation of outcomes gathered is appealing for application in a research environment.[2,3] Combining data from diverse clinical setting also strengthens the generalisability of findings.[3] The approach requires testing the necessary mechanisms to access the data and meeting regulatory requirements.

Non-attendance at paediatric urology outpatient appointments not only results in the patient's failure to receive medical care but also potentially indicates a situation where the patient, parent or caregiver is having trouble negotiating the medical care system. In addition, the no-shows (NS) waste health care resources, and the absence of any communication prevents rescheduling or fitting another patient into the appointment slot to improve clinic efficiency. NS rates range from 15% to 30% of scheduled paediatric and general outpatient appointments/visits.[4,5] Previous studies looking specifically at patient cancellations with general clinic all-ages environments have indicated that there are multiple factors resulting in NS. Clinical factors include difficulty making an appointment; poor patient–staff relationships and difficult communication with the clinic; longer waiting times between scheduling and appointment availability and having a follow-up rather than initial appointment. Demographic factors include being female; living in a lower socio-economic area and having less education.[4–8] Among paediatric urology outpatient surgical cancellations, the majority of cancellations were because of illness, but several preventable causes including financially related issues and fasting violations were also identified.[9] However, there are few data regarding specific paediatric non-attendance risk factors in the outpatient clinic setting.

This study is designed to identify patient demographic variables predictive of patient NS to outpatient appointments while concurrently determining the feasibility of utilising automatically collected data from the Epic EHR to conduct multicentre comparative effectiveness research (CER).

METHODS

After the institutional review board permission was obtained at each of three tertiary paediatric urology practices that utilise the Epic (Madison, WI) EHR system as part of the routine clinical practice (Children's Hospital Colorado (CHCO), Rady Children's Hospital (RCHSD) and University of Virginia Hospital (UVA)), we performed a retrospective review of all urology outpatient appointments scheduled from 1 October 2010 to 30 September 2011. Full implementation of Epic was completed at CHCO in 2006 and at RCHSD and UVA in 2010.

Data were extracted from the EHR at each site using standard automated electronic data extraction techniques. Data acquired included the type of appointment (new or follow-up); primary or satellite urology clinic; type of provider (medical doctor or mid-level provider); date of appointment and days from scheduling to appointment. Age (0–18 years); payer information, patient home zip code and time of appointment were available for CHCO and RCHSD NS patients; institutional regulations prevented extraction at UVA. Time to appointment was calculated as the time between the day the appointment was scheduled and the scheduled date. NS patients were identified as those patients who had an outpatient clinical appointment scheduled during the study period who were marked in the system as a NS. NS is assigned to a patient who does not arrive for a scheduled appointment and has not made contact indicating that he or she would miss the appointment.

Data were de-identified at each site and then merged into a single data set for analysis. Data quality evaluation of key study variables included assessment of age and date ranges as well as examining missing data trends. The consistency of coding terminologies against the Epic reference was also evaluated. SPSS version 21 (IBM, Chicago) was used to test initial associations of demographic and clinical variables using χ^2 for categorical data and analysis of variance for continuous variables. Using identified predictor variables, odds for being a NS appointment were tested using multivariate logistic regression. Cost data, including work revenue value units (wRVUs), were calculated based on clinic and physician averages for new and established patients and using the year's mix for appointment length.

RESULTS

Discrete data were drawn from data captured in the day-to-day care of patients. The available data represent a combined three-site population of 28,715 with 2994 NS patients identified, resulting in a mean NS rate of 10.4%. A common data set of provider type, visit type, clinic (primary or satellite), days from appointment and season was created for the three institutions. A larger data set containing age, sex, payer and zip code was available for CHCO and RCHSD.

Among the three clinics, NS appointments were scheduled for a mean of 53.73 days after requesting an appointment, but varied significantly among institution (F-value 113.65; $p < 0.001$) (Table 1). Clinic level bivariate analysis demonstrated institutional difference among variables of interest. Return patients were significantly associated with NS status at CHCO ($p = 0.03$) and UVA ($p < 0.001$), and trended at RCHSD ($p = 0.08$). Scheduling an appointment at RCHSD during the spring was associated with non-attendance ($p = 0.01$) (Table 2).

Bivariate analysis of the three-institution combined data determined that primary clinic site, rather than satellite, and an appointment with a mid-level provider were associated with non-attendance (Table 2). A multivariate logistic regression model was created to better understand the strength of the factors associated with a NS appointment. The model included institution, provider type, visit type, season and number of scheduled days out. An appointment with mid-level provider (OR 1.70 95% CI (1.56, 1.85)) and increased number of days between scheduling and appointment (15–28 days OR 1.24 CI (1.09, 1.41); 29+ days OR 1.70 CI (1.53, 1.89)) were significantly associated with NS appointments among the three clinics (Table 3). Being a return patient and the location of the visit were not significant in the model.

A multivariate logistic model was created with the CHCO and RCHSD data, which had additional variables of age and payer information (Table 4). In this model, sex, age and season were not significant, but being seen at a primary clinic (OR 2.24 (1.88, 2.66)); government-sponsored insurance (OR 3.41 (3.08, 3.76)) and new patient status (OR 1.20 (1.03, 1.31)) were significant. Mid-level provider (OR 1.71 (1.54, 1.89); increased days to scheduled appointment (15–28 days OR 1.28 (1.11, 1.47); 29+ days OR 1.59 (1.42, 1.78) were also all associated with NS patients within the logistic model.

Table 1 NS characteristics by institution

	CHCO			RCHSD			UVA	
		Mean	*P*-value		Mean	*P*-value	Mean	*P*-value
Days out			*<0.001*[a]			*<0.001*[a]		*<0.001*[a]
NS		34.9			54.6		73.4	
Show		26.8			44.2		55.0	

	N	(%)		N	(%)		Nc	(%)	*P*-value
Visit type			*0.03*[b]			0.08[b]			*<0.001*[b]
New	309	(8.9)		722	(9.6)		97	(9.0)	
Return	314	(10.5)		1008	(8.9)		303	(14.2)	
Clinic			*<0.001*[b]			*<0.001*[b]			*<0.001*[b]
Primary	493	(10.5)		1674	(9.6)		588	(20.9)	
Satellite	130	(7.4)		57	(4.0)		52	(1.8)	
Season			0.49			*0.01*			0.47
Winter	158	(9.9)		426	(9.3)		154	(19.8)	
Spring	162	(9.7)		516	(10.2)		166	(17.0)	
Summer	148	(10.4)		403	(8.5)		192	(19.0)	
Fall	155	(8.8)		386	(8.7)		128	(18.8)	
Provider			*<0.001*[b]			*<0.001*[b]			*<0.001*[b]
Physician	396	(8.5)		1102	(7.6)		285	(16.0)	
Mid-level	227	(12.7)		629	(14.7)		355	(21.2)	

[a]*T*-test
[b]Chi-squared test
[c]Reduced data due to implementation of scheduling module

Table 2 Bivariate analysis combined data

	Not seen	Seen	*P*-value[a]
Visit type			0.08
New	1128	10,900	
Return	1625	14,820	
Clinic			<0.001
Primary	2755	19,916	
Satellite	239	2994	
Sex			0.52
Female	803	7662	
Male	1551	15,248	
Season			0.09
Winter	738	6218	
Spring	844	6857	
Summer	743	62,420	
Fall	669	6226	
Provider type			<0.001
Physician	1783	19,186	
Mid-level	1211	6535	

[a]Chi-squared test

Table 3 Logistic model three institutions

	Adjusted odds ratio	95% CI	*P*-Value
Location			
CHCO (Ref)			
RCHSD	0.90	(0.78, 1.04)	NS
UVA	1.03	(0.91, 1.17)	NS
Clinic location primary clinic	1.03	(0.94, 1.12)	NS
Satellite (Ref)	1.00		
Provider mid-level	1.69	(1.55, 1.84)	<0.001
Physician (Ref)	1.00		
Patient type return	1.20	(1.09, 1.32)	<0.001
New (Ref)	1.00		
Days from making appointment			
0–14 days (Ref)	1.00		
15–28 days	1.24	(1.09, 1.41)	0.001
29+ days	1.70	(1.53, 1.89)	<0.001

Table 4 CHCO and RCHSD logistic model

	Adjusted odds ratio	95% CI	*P*-Value
Clinic location			
Primary clinic	2.24	(1.88, 2.67)	<0.001
Satellite provider	1.00		
Mid-level	1.71	(1.54, 1.89)	<0.001
Physician appointment	1.00		
Return	1.20	(1.09, 1.32)	<0.001
New Insurance	1.00		
Government payer	3.40	(3.09, 3.76)	<0.001
Private	1.00		
Age group			
Age 0–5 (Ref)	1.00		
Age 6–12	1.10	(0.99, 1.22)	NS
Age 13–18	1.10	(0.96, 1.26)	NS
Days from making appointment			
0–14 days (Ref)	1.00		
15–28 days	1.28	(1.11, 1.47)	0.001
29+ days	1.59	(1.42, 1.78)	<0.001

The addition of payer type in this section of the study population had a significant effect on the significance of the association among the variables of interest.

Using a historic clinic mix of 15- and 30-min appointments, it was estimated that the NS appointments represented more than 920 h of clinic time, or 4510 equivalent wRVUs, that were lost at the three clinics. Estimating $52.50 physician compensation per wRVU, lost revenue approaches $237,000.

DISCUSSION

Principal findings

Among a small network of paediatric Epic EHR users, we demonstrated that we could perform the necessary extraction data for a multi-site collaborative research project within a newly implemented EHR environment, which was an important first step. We used broad inclusion criteria for the study population that were purposely wide ranging and designed to capture a diverse sample. We were quickly able de-identify at a level that met needed regulations but still permitted analysis. As a result, we have a proof-of-concept study conducted at three disparate geographic locations with heterogeneous populations and different referral patterns. It was feasible, practical and expedient to use CER to measure reasons for non-attendance in a natural practice environment rather than a controlled setting.

Comparison with the literature

The Agency for Healthcare Research Quality has defined CER as research designed to inform health care decisions by providing evidence on effectiveness, benefits and harms of different treatment options. Evidence is generated from research studies that compare drugs, medical devices, tests, surgeries or ways to deliver health care. Information resulting from CER can help to identify interventions that are effective under various circumstances for patients, providers and policy makers.[10–12] As such, there has been substantial federal investment in CER with the goal of studying populations and outcomes of real-world clinical practices.[13] Using this framework, it was possible to address the anecdotal sense among clinicians that there were patient characteristics associated with appointment NS and use the discrete variables that are already part of EHR obtained through automated data extraction to test this question.

Previous studies conducted on adults or populations of all ages using survey and interview methodology have found that site of care is a powerful predictor of not attending an appointment. Patients also miss appointments according to the hour of the appointment, the source of referral, delays in the appointment date and if a doctor is more junior.[8,14] Using a novel research data gathering approach, we found similarities within our paediatric patients and with previous adult and general clinic studies. We were also able to efficiently and effectively attach a cost estimate to the situation using the extracted data. Using an accurate estimate of NS, the year's mix for appointment length and the comparable wRVUs, the costs as well as the clinical issues can be logically addressed. Strengths of our study include the large sample size and demonstrated interoperability of individually installed Epic systems across three geographically disparate institutions.

Limitations of the method

It has been noted that issues can arise when trying to combine data with different storage structures, those using dissimilar ontologies or those from incompatible systems.[2] While all three institutions were using Epic, they were in different phases of implementation, and one of the sites was still using a combination of the old scheduling system and Epic, which resulted in not having some variables from one of the sites. It has been well documented that working with data collected for clinical and billing uses rather than research often lacks the necessary temporal relationship to determine disease cause or to define a diagnosis.[13] Several demographic and environmental variables of interest were incomplete or not available at this time. We did not measure race, language, other surrogates for socio-economic status or parent/caregiver educational level, which have demonstrated an association with non-attendance.[14] Another issue encountered is establishing patterns of non-attendance.[2] Neal et al.[7] found within a primary care population in the United Kingdom that 90% of the non-attenders subsequently were seen for an appointment within three months. We would need to employ more complex algorithms and have access to more confirmatory personal health information to perform similar analysis. Since many of the NS referral patients had never been part of the EHR systems, we could not rely on auxiliary techniques, such as capturing race based on visual ascertainment or capturing language using information about an interpreter requested or the language in which forms were completed as other studies have done.[14] There were limitations in trying to identify the diagnosis or diagnoses associated with the reason for the appointment since assigning a diagnosis or entering a condition is normally part of the workflow during a patient encounter, and the patients had not actually been seen for their requested appointments. Not all challenges were technical. In compliance with Health Insurance Portability and Accountability Act and individual institutional regulations, we were restricted in the extraction of some demographic information such as age and gender. We found the regulatory models regarding such data and sharing of data to be too restrictive for such a multi-institutional project, even though data were de-identified at the site level.

Call for further research

Data quality issues can arise because of variation in data capture, local business rules and clinical workflows.[3,15] We noted that differences in such flow resulted in some variables of interest being available in two of three institutions. Frequently, health care providers are not trained to gather data for research purposes, and the EHR is not designed for research. The amount of missing or incomplete data can reinforce the concept that the EHR may contain inaccurate data.[2,13,16] The fact that payer had such a significant association and resulted in a significant change in statistical association among several variables in the CHCO/RCHSD model points to the need for a more comprehensive data pull with access to additional data.

While obtaining descriptive and multivariate statistics was an important focus at this level of the intervention, multi-institutional data extraction feasibility was also an important consideration. This project evaluated the type and robustness of data availability. While we generally found adequate scientific content and data access, we found that additional approaches will be needed for future prospective data collection to describe the medical utilization within research populations of interest.

Overcoming the limitation of working with available data requires extra work on the part of the researchers. Recognizing that data not gathered specifically for research purposes may be incomplete and unreliable does not make the data unusable, but it does constrain the uses to which the data can be put and the inferences that can be drawn from them.[1,2] The overall scope of our project was adjusted to answer questions using data that were available. We framed our research questions so that they could analyze the much more easily extracted structured and discrete data rather than narrative text.

Patients may not understand what happens in the clinic when there is a NS appointment.[17] The technology is available in Epic through the patient portal MyChart to capture patient report electronically and coordinate with the medical record.[18] A needed next step to continue to understand and remediate patient non-attendance is to include patient-reported outcomes and patient feedback regarding their reasons for non-attendance.

CONCLUSIONS

We were able to demonstrate interoperability among the EHRs and obtain data to identify rapidly, efficiently and effectively sub-populations for 1) intervention, 2) further study and 3) more complex analysis. Institution-specific practices may have an underlying influence on resulting data availability and collection, and will be the goal for future collaborative research efforts. It is possible to use information from a paediatric EHR to perform CER that can inform health care decisions regarding roadblocks or factors influencing non-compliance.

ACKNOWLEDGEMENT

Dr. Bush's time was supported in part by grant number K99 HS022404-01 from the Agency for Healthcare Research and Quality.

REFERENCES

1. Wasserman RC. Electronic medical records (EMRs), epidemiology, and epistemology: reflections on EMRs and future pediatric clinical research. *Academic Pediatrics* 2011;11(4):280–7. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3138824&tool=pmcentrez&rendertype=abstract. http://dx.doi.org/10.1016/j.acap.2011.02.007.
2. Embi PJ, Hebert C, Gordillo G, Kelleher K and Payne PRO. Knowledge management and informatics considerations for comparative effectiveness research: a case-driven exploration. *Medical Care* 2013 Aug;51(8 Suppl 3):S38–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23793050. http://dx.doi.org/10.1097/MLR.0b013e31829b1de1.
3. Brown JS, Kahn M and Toh S. Data quality assessment for comparative effectiveness research in distributed data networks. *Medical Care* 2013 Aug;51(8 Suppl 3):S22–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23793049. http://dx.doi.org/10.1097/MLR.0b013e31829b1e2c.
4. Neal RD, Lawlor DA, Allgar V, Colledge M, Ali S, Hassey A et al. Missed appointments in general practice: retrospective data analysis from four practices. *The British Journal of General Practice* 2001 Oct;51(471):830–2. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1314130&tool=pmcentrez&rendertype=abstract.
5. Ross LV, Friman PC and Christophersen ER. An appointment-keeping improvement package for outpatient pediatrics: systematic replication and component analysis. *Journal of Applied Behavior Analysis* 1993 Jan;26(4):461–7. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid= 1297871&tool=pmcentrez&rendertype=abstract. http://dx.doi.org/10.1901/jaba.1993.26-461.
6. Martin C, Perfect T and Mantle G. Non-attendance in primary care: the views of patients and practices on its causes, impact and solutions. *Family Practice* 2005 Dec;22(6):638–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16055472.
7. Neal RD, Hussain-Gambles M, Allgar VL, Lawlor DA and Dempsey O. Reasons for and consequences of missed appointments in general practice in the UK: questionnaire survey and prospective review of medical records. *BMC Family Practice* 2005 Nov;6:47. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1291364&tool=pmcentrez&rendertype=abstract http://dx.doi.org/10.118/1471-2296-6-47.
8. Perron NJ, Dao MD, Kossovsky MP, Miserez V, Chuard C, Calmy A et al. Reduction of missed appointments at an urban primary care clinic: a randomised controlled study. *BMC Family Practice* 2010 Jan;11:79. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2984453&tool=pmcentrez&rendertype=abstract. http://dx.doi.org/10.1186/1471-2296-11-79.
9. Pohlman GD, Staulcup SJ, Masterson RM and Vemulakonda VM. Contributing factors for cancellations of outpatient pediatric urology procedures: single center experience. *Journal of Urology* 2012 Oct;188(4 Suppl):1634–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22910272. http://dx.doi.org/10.1016/j.juro.2012.03.
10. Gallego B, Dunn AG and Coiera E. Role of electronic health records in comparative effectiveness research. *Journal of Comparative Effectiveness Research* 2013 Nov;2(6):529–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24236790 . http://dx.doi.org/10.2217/cer.13.65.
11. Miriovsky BJ, Shulman LN and Abernethy AP. Importance of health information technology, electronic health records, and continuously aggregating data to comparative effectiveness research and learning health care. *Journal of Clinical Oncology* 2012 Dec;30(34):4243–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23071233. http://dx.doi.org/10.1200/JCO.2012.42.8011.
12. Bayley KB, Belnap T, Savitz L, Masica AL, Shah N and Fleming NS. Challenges in using electronic health record data for CER: experience of 4 learning organizations and solutions applied. *Medical Care* 2013 Aug;51(8 Suppl 3):S80–6. Available from: http://www.ncbi.nlm.nih.gov/pubme/23774512. http://dx.doi.org/10.1097/MLR.0b013e31829b1d48.
13. Hersh WR, Weiner MG, Embi PJ, Logan JR, Payne PRO, Bernstam E V et al. Caveats for the use of operational electronic health record data in comparative effectiveness research. *Medical Care* 2013 Aug;51(8 Suppl 3):S30–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23774517. http://dx.doi.org/10.1097/MLR.0b013e31829b1dbd.
14. Lasser KE, Mintzer IL, Lambert A, Cabral H and Bor DH. Missed appointment rates in primary care: the importance of site of care.

Journal of Health Care for the Poor and Underserved 2005 Aug;16(3):475–86. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16118837.

15. Liaw S-T, Taggart J, Yu H and de Lusignan S. Data extraction from electronic health records – existing tools may be unreliable and potentially unsafe. *Australian Family Physician* 2013 Nov;42(11):820–3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24217107.

16. Curcin V, Soljak M and Majeed A. Managing and exploiting routinely collected NHS data for research. *Informatics in Primary Care* 2012 Jan;20(4):225–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23890333. http://dx.doi.org/.

17. Lacy NL, Paulman A, Reuter MD and Lovejoy B. Why we don't come: patient perceptions on no-shows. *The Annals of Family Medicine* 2004;2(6):541–5. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi? artid=1466756&tool=pmcentrez&rendertype=abstract. http://dx.doi.org/10.1370/afm.123.

18. Wu AW, Kharrazi H, Boulware LE and Snyder CF. Measure once, cut twice–adding patient-reported outcome measures to the electronic health record for comparative effectiveness research. *Journal of Clinical Epidemiology* 2013 Aug;66(8 Suppl):S12–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23849145. http://dx.doi.org/10.1016/j.jclinepi.2013.04.005.