

# Using the Electronic Health Record to Identify Educational Gaps for Internal Medicine Interns

Alvin Rajkomar, MD  
Sumant R. Ranji, MD  
Bradley Sharpe, MD

## ABSTRACT

**Background** An important component of internal medicine residency is clinical immersion in core rotations to expose first-year residents to common diagnoses.

**Objective** Quantify intern experience with common diagnoses through clinical documentation in an electronic health record.

**Methods** We analyzed all clinical notes written by postgraduate year (PGY) 1, PGY-2, and PGY-3 residents on medicine service at an academic medical center July 1, 2012, through June 30, 2014. We quantified the number of notes written by PGY-1s at 1 of 3 hospitals where they rotate, by the number of notes written about patients with a specific principal billing diagnosis, which we defined as *diagnosis-days*. We used the International Classification of Diseases 9 (ICD-9) and the Clinical Classification Software (CCS) to group the diagnoses.

**Results** We analyzed 53 066 clinical notes covering 10 022 hospitalizations with 1436 different ICD-9 diagnoses spanning 217 CCS diagnostic categories. The 10 most common ICD-9 diagnoses accounted for 23% of diagnosis-days, while the 10 most common CCS groupings accounted for more than 40% of the diagnosis-days. Of 122 PGY-1s, 107 (88%) spent at least 2 months on the service, and 3% were exposed to all of the top 10 ICD-9 diagnoses, while 31% had experience with fewer than 5 of the top 10 diagnoses. In addition, 17% of PGY-1s saw all top 10 CCS diagnoses, and 5% had exposure to fewer than 5 CCS diagnoses.

**Conclusions** Automated detection of clinical experience may help programs review inpatient clinical experiences of PGY-1s.

## Introduction

In nonprocedural specialties, quantifying and tracking clinical experiences of residents is challenging. Recording the principal diagnoses of patients managed by residents is feasible with a modern electronic health record (EHR), but existing studies using this approach have had limitations.<sup>1-3</sup> Prior studies have attributed clinical experience to the final clinician who saw a patient, which may not capture the experience of multiple other physicians who could have admitted or followed the daily progress of patients.<sup>2</sup> Accurately capturing clinical diagnoses seen by trainees is important but difficult, as granular billing codes may be too specific to quantify clinical experience.<sup>1</sup> In addition, relying on residents themselves to input diagnostic clusters<sup>3</sup> imposes additional documentation burdens, and may not provide accurate data.

We developed a method to characterize inpatient clinical experience for internal medicine postgraduate year (PGY) 1 trainees by analyzing the diagnoses associated with their clinical notes, and by quantifying their experience by the number of notes written about patients with specific diagnoses. We also

grouped relevant diagnoses together according to standard clinical classification diagnoses; this allowed separate, but closely related, International Classification of Diseases 9 (ICD-9) codes to count equally to a diagnostic category.

## Methods

We studied clinical notes written by internal medicine residents at the University of California, San Francisco (UCSF) from July 1, 2012, through June 30, 2014. The internal medicine residency program comprises 180 trainees, including 62 PGY-1s. We analyzed data from the UCSF Medical Center, 1 of 3 hospitals through which trainees rotate. The PGY-1s generally spend 2 to 3 months on inpatient medicine rotations at that site, and are responsible for writing daily notes on their patients. On occasional days, PGY-2 and PGY-3 residents may also write notes. We analyzed only the experiences of the PGY-1s who spent at least 2 months on an internal medicine rotation, identifying history and physical examination, progress, and discharge summary notes in the EHR system (Epic, Verona, WI) for patients who were discharged from the inpatient medicine service.

The outcome of the study was the number of notes written about a specific diagnosis. For each note, we

DOI: <http://dx.doi.org/10.4300/JGME-D-16-00272.1>

collected the single principal diagnosis ICD-9 code associated with that hospitalization. We obtained the principal diagnosis from billing data after discharge, and mapped multiple individual ICD-9 diagnosis codes into diagnostic groups according to categories defined by the Clinical Classifications Software (CCS; Agency for Healthcare Research and Quality, Rockville, MD).

This study was approved by the UCSF Institutional Review Board.

We calculated the most frequent diagnoses documented using 2 methods. First, we tabulated clinical experiences for PGY-1s if they had any note type associated with specific ICD-9 discharge diagnoses. For example, if a patient with “unspecified septicemia” had notes written by 2 PGY-1s at different times, both would get credit for “experience,” even if 1 of them only authored the discharge summary or a single progress note. For study simplicity, we used the terminology *seeing* or *caring for* a patient to indicate writing a note about that patient.

For the second analysis, we represented clinical experience through *diagnosis-days*, which were the number of inpatient days spent caring for patients with each diagnosis. For example, for the above patient with “septicemia,” if 1 PGY-1 cared for the patient for 5 days and another resident only for 1 day, the first would have 5 diagnosis-days of experience with that CCS code, while the second would have only 1 day of experience. Descriptive statistics were used to demonstrate the experience of PGY-1s with different diagnostic categories. R software version 3.1.1 (The R Foundation for Statistical Computing, Vienna, Austria) was used for data preparation and for descriptive analysis. Data extraction and analysis were performed by 1 author (A.R.), who has data science training and is certified to perform Epic data queries. One author (A.R.) also spent 20 hours to determine how to optimally extract the data from the EHR, an estimated 10 hours to extract and validate the data, and 5 hours to perform the analysis.

## Results

There were 10 022 hospitalizations during the study period, generating 53 066 notes written by residents: 8852 histories and physical examination reports, 35 115 progress notes, and 9099 discharge summaries. Because some patients were admitted by attending physicians and some were transferred from other services, there were more discharge summaries than there were histories and physicals. There were 1436

### What was known and gap

Medical programs want better ways to assess the clinical experience of their residents.

### What is new

A study analyzed internal medicine interns' clinical notes in the electronic health record, quantifying their exposure to common diagnoses using the International Classification of Diseases 9 (ICD-9).

### Limitations

Use of principal billing diagnoses may not fully reflect interns' experience.

### Bottom line

Of interns who spent 2 months on the inpatient service, 3% were exposed to all of the top 10 ICD-9 diagnoses, whereas 31% had experience with fewer than 5 diagnoses.

different billing diagnoses spanning 217 CCS diagnostic categories.

During the study period, 122 internal medicine PGY-1s on hospital medicine rotations authored notes. We limited analysis to the 107 (88%) internal medicine PGY-1s who spent a median of 2.5 months on the hospital medicine service (interquartile range [IQR] 2–2.5). These PGY-1s wrote a total of 23 718 notes, averaging 221.7 notes per individual.

### Top ICD-9 Diagnoses

The 10 most common ICD-9 discharge diagnoses accounted for 22% of hospitalizations and 23% of diagnosis-days. They were unspecified septicemia, pneumonia, acute kidney failure, urinary tract infection, obstructive chronic bronchitis with exacerbation, pneumonitis due to inhalation of food or vomitus, human immunodeficiency virus (HIV) disease, acute pancreatitis, encounter for antineoplastic chemotherapy, and septicemia due to *escherichia coli*.

Each PGY-1 saw a median of 51 ICD-9 diagnoses (IQR 41–61.5). Three of 107 (3%) saw all of the top 10 ICD-9 diagnoses during their rotations, and 12 (11%) saw 9 of the 10. In contrast, 104 (97%) did not see a patient with at least 1 of the top 10 diagnoses, and 33 (31%) had experience with fewer than 5 of the top 10 diagnoses.

### Top CCS Diagnoses

The 10 most commonly encountered CCS groupings by diagnosis-days (not just by diagnosis) accounted for 37% of hospitalizations and 40% of diagnosis-days. They included septicemia (except in labor); complications from devices, implants, or grafts; secondary malignancies; pneumonia (except that

**TABLE**  
PGY-1s Who Saw at Least 1 of the Top 10 Most Common Diagnoses (N = 107)<sup>a</sup>

Diagnosis	PGY-1s Who Wrote Note About Diagnosis, <sup>b</sup> n (%)
Septicemia (except in labor)	107 (100)
Complication of device, implant, or graft <sup>c</sup>	99 (93)
Secondary malignancies	98 (92)
Pneumonia (except that caused by tuberculosis or sexually transmitted disease)	94 (88)
Gastrointestinal hemorrhage	90 (84)
Acute and unspecified renal failure	85 (79)
Alcohol-related disorders	80 (75)
Respiratory failure, insufficiency, or arrest (adult)	67 (63)
Diabetes mellitus with complications	80 (75)
Urinary tract infections	77 (72)

<sup>a</sup> The diagnosis column indicates the top 10 most common Clinical Classification Software diagnoses, as ranked by the number of “diagnosis-days” documented by residents.

<sup>b</sup> The right-hand column indicates the percentage of PGY-1s in a 2-year period who wrote any clinical note about a hospitalization with the corresponding diagnosis.

<sup>c</sup> Complications of devices, implants, or grafts include central line infection; lung transplant complications; kidney transplant complications; complications of medical care, not elsewhere classified; dialysis access complications; infected joint prostheses; urinary tract infections; and other disorders of intestines.

caused by tuberculosis or sexually transmitted diseases); gastrointestinal hemorrhage; acute and unspecified renal failure; alcohol-related disorders; respiratory failure, insufficiency, or arrest (adult); diabetes mellitus with complications; and urinary tract infections. We report the percentage of PGY-1s who wrote a note corresponding to 1 of these diagnoses in the TABLE.

Each PGY-1 saw a median of 38 different diagnostic categories (IQR 32–44), with 18 of 107 (17%) seeing all of the top 10 CCS diagnoses during their internship, and 55 (51%) who saw 9 of the 10. In contrast, 89 (83%) did not see a patient with at least 1 of the top 10 diagnoses, and 5 (5%) saw fewer than 5 of the top 10 CCS diagnoses.

## Discussion

We used EHR notes to quantify PGY-1 exposure to principal billing diagnoses in an inpatient medical service at a single hospital. We found variations in exposure, as defined by 2 measures: the number of unique ICD-9 diagnoses and the number of diagnosis-days of CCS groupings. This innovative approach to quantifying clinical exposure did not require a significant amount of time from a clinician data scientist.

There are 2 implications from these results for graduate medical education practice. First, the use of progress notes and billing codes generated as part of clinical practice may allow programs to track resident clinical experiences without extensive addi-

tional data gathering. Second, we found that using diagnosis-days of CCS groupings may allow programs to better characterize clinical experience by grouping together relevant ICD-9 codes and by reflecting the duration of the experience. Graduate medical education program leadership could provide the results of an analysis of clinical experience directly to PGY-1s to help them identify gaps in their training and to subsequently seek out specific clinical experiences. Programs could also tailor curricular experiences or clinical schedules to meet the needs of individual residents. Our study should be feasible for other institutions to replicate. Our institution uses a widely deployed EHR, and although a physician informaticist performed our data extraction and analysis, the data can be extracted from standard metadata about clinical notes, such as the note author and entry date.

Our study has several limitations. First, we only used the principal coded billing diagnosis, which excluded secondary diagnoses and may not reflect the major clinical experience of a hospitalization (eg, treating delirium in a patient with pneumonia). Second, PGY-1s could have obtained experience from a patient with a particular diagnosis without writing a note (eg, by attending bedside rounds) or during nonhospitalization encounters (eg, emergency department and outpatient visits), and our methodology may underestimate “experience.” Finally, we were only able to study notes at 1 of the 3 clinical sites where residents rotate.

## Conclusion

It is feasible to use EHR data to quantify the degree of exposure to primary diagnoses among residents in order to potentially identify gaps in clinical experience. Future research should explore how to more precisely quantify clinical experiences using EHR data, and how to use such data to promote optimal educational experiences.

## References

1. Bachur RG, Nagler J. Use of an automated electronic case log to assess fellowship training: tracking the pediatric emergency medicine experience. *Pediatr Emerg Care.* 2008;24(2):75–82.
2. Nagler J, Harper MB, Bachur RG. An automated electronic case log: using electronic information systems to assess training in emergency medicine. *Acad Emerg Med.* 2006;13(7):733–739.
3. Sequist TD, Singh S, Pereira AG, et al. Use of an electronic medical record to profile the continuity clinic

experiences of primary care residents. *Acad Med.* 2005;80(4):390–394.



**Alvin Rajkomar, MD**, is Assistant Professor, Department of Medicine, University of California, San Francisco; **Sumant R. Ranji, MD**, is Professor, Department of Medicine, Zuckerberg San Francisco General Hospital and University of California, San Francisco; and **Bradley Sharpe, MD**, is Professor, Department of Medicine, University of California, San Francisco.

Funding: The authors report no external funding source for this study.

Conflict of interest: The authors declare they have no competing interests.

These results were presented as a poster at the Society of Hospital Medicine Annual Meeting, Las Vegas, Nevada, March 24–27, 2014.

Corresponding author: Alvin Rajkomar, MD, University of California, San Francisco, 533 Parnassus Avenue, Suite 127A, San Francisco, CA 94143-0131, 415.476.1528, fax 415.502.1976, alvin.rajkomar@ucsf.edu

Received May 3, 2016; revisions received July 11, 2016, and August 16, 2016; accepted September 30, 2016.