Perspective

Metrics for assessing physician activity using electronic health record log data

Christine A. Sinsky, Adam Rule, Genna Cohen, Brian G. Arndt, Tait D. Shanafelt, Christopher D. Sharp, Sally L. Baxter, Ming Tai-Seale, Sherry Yan, You Chen, Julia Adler-Milstein, and Michelle Hribar

1Department of Medicine, American Medical Association, Chicago, Illinois, USA, 2Department of Medical Informatics and Clinical Epidemiology, Oregon Health Sciences University, Oregon, USA, 3Department of Medicine, Mathematica, Washington, DC, USA, 4Department of Family Medicine and Community Health, School of Medicine and Public Health, University of Wisconsin, Madison, Wisconsin, USA, 5Division of Hematology, Department of Medicine, Stanford University, Stanford, California, USA, 6Division of General Internal Medicine, Department of Medicine, Stanford University, Stanford, California, USA, 7Department of Biomedical Informatics, University of California, San Diego, San Diego, California, USA, 8Viterbi Family Department of Ophthalmology, Shiley Eye Institute, University of California, San Diego, San Diego, California, USA, 9Department of Family Medicine and Public Health, University of California, San Diego, San Diego, California, USA, 10Department of Medicine, Sutter Health, Walnut Creek, California, USA, 11Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, USA, and 12Department of Medicine, University of California, San Francisco, San Francisco, California, USA

Corresponding Author: Christine A. Sinsky, MD, 330 N. Wabash Ave, Suite 39300, Chicago, IL 60611-5885, USA; christine.sinsky@ama-assn.org

Received 4 November 2019; Revised 10 December 2019; Editorial Decision 15 December 2019; Accepted 17 December 2019

ABSTRACT

Electronic health record (EHR) log data have shown promise in measuring physician time spent on clinical activities, contributing to deeper understanding and further optimization of the clinical environment. In this article, we propose 7 core measures of EHR use that reflect multiple dimensions of practice efficiency: total EHR time, work outside of work, time on documentation, time on prescriptions, inbox time, teamwork for orders, and an aspirational measure for the amount of undivided attention patients receive from their physicians during an encounter, undivided attention. We also illustrate sample use cases for these measures for multiple stakeholders. Finally, standardization of EHR log data measure specifications, as outlined here, will foster cross-study synthesis and comparative research.

Key words: metric, operational efficiency, EHR log data, time studies, burnout

INTRODUCTION

Electronic health records (EHRs) have transformed the daily work of clinicians. Widely implemented over the past 2 decades, with an investment of approximately $36 billion in federal funding, EHRs are now ubiquitous in ambulatory (86%) and inpatient (96%) settings. Despite this investment and widespread adoption, the evidence for the impact of EHRs on quality, cost, and efficiency is mixed, and there is growing evidence of a negative impact on the work lives of physicians. In particular, the use of an EHR has been associated with higher rates of burnout. A direct observation time-motion study in 2016 found physicians in the ambulatory setting spend half of the workday on the EHR, requiring nearly 2 hours of EHR and desk work for every 1 hour of direct clinical face time with patients. Subsequent studies using EHR log data similarly found that approximately half of the physician workday is spent on the EHR. In addition, physicians may allocate 1-2 hours of personal time each night to EHR tasks. Special
concerns have been raised about the time costs of several clinical tasks, including inbox management, computerized physician order entry, algorithm-generated alerts, visit-note documentation, and the overall volume of administrative burden mediated through the EHR.

While EHR log data has shown promise in measuring time spent on these clinical activities, the use of EHR log data to further understand the clinical environment is a nascent science. All certified EHRs in the United States are required to maintain EHR logs, tracking who accesses which patient records at what time, and what they do in the record, such as view, edit, or delete patient information. While originally designed to support auditing of inappropriate record access, these logs provide an opportunity to study clinical activities unobtrusively and at scale, including how time is distributed among various tasks and role types, and how interventions, such as new policies, staffing models, or workflows, impact the time spent within the EHR. A systematic review identified 85 studies using EHR log data. Most did not detail their definitions of EHR task or methodology in estimating EHR time, nor did they validate their measures. Some researchers accessed raw data files from EHR logs, using extensive data cleaning and preprocessing to compute measures such as total EHR time. Others used “off-the-shelf” processed data provided by vendors, typically generated by proprietary algorithms and shaped by the assumptions and definitions which are not always known to the researcher or reader. This variation hinders cross-study synthesis and comparative effectiveness research.

Therefore, a network of researchers and other stakeholders working with EHR log data collaborated to standardize EHR audit log measures that capture different dimensions of the physician EHR work. In this commentary, we propose 7 core EHR use measures, identify potential use cases for these measures across multiple stakeholders, highlight unresolved methodologic considerations, and address future directions for research and use. This work builds on prior work outlining potential measures to evaluate the practice environment by refining and expanding the core measure set and providing detailed specification.

**FACTORS GUIDING MEASURE SELECTION**

We worked with multiple stakeholders to select measures that are feasible and relevant to clinical and operational decision making. We identified time as an important unit of measure, as it is a commodity in limited supply for physicians and other health professionals. We also recognized the need to normalize time-based data to allow comparisons across different degrees of clinical effort (i.e., part time vs full time); we chose an 8-hour period of scheduled patient time as the unit of normalization. In designing these measures, we prioritized simplicity and feasibility over complexity and granularity, recognizing that there will be variations in clinical circumstances that are not fully captured with these initial metrics. In addition, while time alone does not capture all dimensions of work (e.g., the content of work and appropriateness of tasks to different levels of training or certification), time-based metrics are relevant, interpretable, and form a foundation from which to expand upon in the future. We also recognize that these measures are best suited to application in the ambulatory setting and will need to be modified for use in the inpatient arena. Finally, while we have focused on physicians, future studies examining EHR use measures for additional roles (medical assistants, pharmacists, nurses, advance practice providers, etc.) will also be beneficial.

**PROPOSED CORE EHR USE MEASURES**

We propose 7 core measures of EHR use which reflect multiple dimensions of practice efficiency (Table 1): total EHR time (EHR-Time); work outside of work (WOW), often referred to as “pjama time”; time on documentation (Note-Time); time on prescriptions (Script-Time); inbox time (IB-Time); teamwork for orders (TWORD); and an aspirational measure for the amount of undivided attention patients receive from their physicians during an encounter, undivided attention (ATTN).

**EXAMPLES OF THE UTILITY OF EHR USE MEASURES**

EHR use measures have multiple uses within research and clinic administration (Table 2). First, EHR use measures can be used to assess the impact of new staffing models, such as advanced team-based care with in-room support. For example, a physician practicing with a strong, skilled team may be able to spend the majority of her time providing undivided attention to her patients, while her empowered staff enters orders, completes the billing invoice, drafts the preliminary visit note, and manages the team’s inbox. Another physician, practicing in a more “the doctor does it all” environment, performs the majority of data entry and inbox management on her own. EHR use measures can help to quantify this difference by establishing the degree to which WOW and Note-Time decrease and TWORD increases when moving to an advanced team-based model of care.

EHR use measures can also inform expectations for patient-scheduled hours for a full-time equivalent position within a given specialty and clinical support environment. A physician sharing a single medical assistant with another physician may be able to accommodate 20 hours of patient-scheduled time per week, because of an additional 20 hours of WOW per week. In another practice within the same specialty, a physician dedicated 2 medical assistants may be able to accommodate 30 hours of patient-scheduled time per week because of a reduction in WOW to 10 h/wk as a result of greater support.

EHR use measures can be a component of predictive analytics. For example, with longitudinal tracking, health systems may find that WOW, IB-Time, and TWORD predict physician risk for distress, burnout, and intent to cut back on clinical effort or leave the practice. EHR use measures can likewise illuminate the impact of regulatory and compliance decisions upon clinician work. For example, will WOW and Note-Time change as a result of the anticipated CPT 2021 modification of documentation guidelines and codes, which eliminates history and exam from criteria in billing office visit codes?

Finally, EHR use measures provide discrete metrics that can be used by multiple stakeholders to promote EHR design improvements to facilitate efficiency of use. These stakeholders include EHR vendors, those responsible for EHR implementation, and those who integrate elements of the EHR designed for regulatory or billing purposes. EHR use measures could help inform understanding of efficient practices and improve future design.

**CHALLENGES AND LIMITATIONS**

Several definitional and methodologic considerations with the use of EHR log data remain unsettled, including those related to measure implementation, validity, normalization, and generalizability.
Table 1. Proposed core EHR use measures

| Measure                        | Abbreviation | Definition and example                                                                 |
|--------------------------------|--------------|----------------------------------------------------------------------------------------|
| Total EHR time                 | EHR-Time₈   | Total time on EHR (during and outside of clinic sessions) per 8 h of patient scheduled time.  
Example: A physician with 32 patient-scheduled hours per week, 20 h of EHR time during scheduled hours, 10 h of WOW each week would have EHR-Time₈ of 30/32 × 8 = 7.5. |
| Work outside of work           | WOW₈        | Time on EHR outside of scheduled patient hours per 8 h of patient scheduled time.  
Example: A physician with 32 scheduled patient hours per week and a total of 10 h of EHR time outside of these scheduled hours, would have WOW₈ = 10/32 × 8 = 2.5. |
| Time on encounter note documentation | Note-Time₈ | Hours on documentation (note writing) per 8 h of scheduled patient time  
Example: A physician with 32 scheduled patient hours per week and a total of 20 h of documentation time (both in the room with the patient and outside of the room) per week would have DocTime₈ of 20/32 × 8 = 5.0. |
| Time on prescriptions          | Script-Time₈| Total time on prescriptions per 8 h of patient scheduled time  
Example: A physician spends 3 h per week on prescription work and has 24 h of scheduled patient time per week. Script₈ = 3/24 × 8 = 1 |
| Time on inbox                  | IB-Time₈   | Total time on inbox per 8 h of patient scheduled time  
Example: A physician spends 10 h per week on Inbox work and has 20 h per week of patient scheduled time. IB₈ = 10/20 × 8 = 4 |
| Teamwork for orders            | TW ORD     | The percentage of orders with team contribution  
Example: A physician working with a team that is empowered to pend, send orders by protocol, or operationalize verbal orders, may compose 25% of the orders from start to finish on their own, while the rest are pended or completed by team members for the physician’s co-signature. In this case, TW_ORD = 75%. |
| Undivided attention            | ATTN       | The amount of undivided attention patients receive from their physician. It is approximated by [(total time per session) minus (EHR time per session)]/total time per session.  
Example: A physician who is actively on the EHR 3 h of a 4-h clinic session would have a lower ATTN score (4-3)/4 = 0.25 than would a physician who was actively on the EHR 1 h of a 4-h clinic session. (4-1)/4 = 0.75. |

EHR: electronic health record.

*For consistency, and to avoid distortion owing to different session lengths, we define work outside of work precisely as that time outside of scheduled patient hours and do not include any “shoulder time” before or after clinic.

Table 2. Sample use cases for EHR log data

| Stakeholder | Objective | Examples of research questions |
|-------------|-----------|--------------------------------|
| Health system leaders | Improve the patient experience | How do increased staffing ratios impact WOW₈, Note-Time₈, and the amount of direct face time patients receive from their physician (ATTN)? How do decreases in WOW₈, Note-Time₈, and IB-Time₈ impact appointment availability and continuity? How do staffing ratios impact the total workload for physicians (ie, patient scheduled time + WOW)? |
| Improve access for patients | Improve professional fulfillment | Can EHR use patterns predict future risk of burnout, other dimensions of distress, or professional fulfillment? |
| Improve professional fulfillment | Improve retention | Can EHR use patterns predict future risk of leaving the organization? |
| Improve retention | Improve recruitment | Physicians may choose to incorporate data regarding work outside of work and total EHR time when evaluating a potential employment offer. |
| Operations leaders | Improve efficiency | Does empowered inbox management by team members reduce physician/APP EHR time after hours? Does advanced team-based care with in-room support reduce work outside of work? |
| Standardize FTE expectations | Understand the costs of compliance decisions | How many patient contact hours is associated with a 40-h total work week in across different specialties? How do changes in staffing ratio, team stability, and team skill level impact work outside of work? How do requirements that only physicians do certain tasks such as order entry, medication reconciliation, and prohibitions against verbal orders impact WOW? |
| Regulators | Understand and evaluate trade-offs involved in a given policy under consideration | Pilot test policies under consideration and evaluate their impact on EHR use metrics before finalizing and implementing. How do requirements that only physicians do certain tasks such as order entry, medication reconciliation and prohibitions against verbal orders impact WOW? What are the time costs associated with prohibiting verbal orders? What are the time costs of requiring 2-factor authentication and password revalidation for nonscheduled prescriptions? |
| Technology vendors | Improve usability | How does an electronic workflow for order entry impact time on orders? |
| Help clients measure and improve practice efficiency | | Provide insight and guidance to help organizations improve efficiency, patient satisfaction, and professional fulfillment for healthcare professionals |

APP: advanced practice provider; ATTN : undivided attention; EHR: electronic health record; FTE: full-time equivalent; IB-Time₈: time on inbox per 8 h of patient scheduled time; Note-Time₈: time on encounter documentation per 8 h of patient scheduled time; WOW₈: work outside of work per 8 h of patient scheduled time.
IMPLEMENTATION

1. Definition of EHR use time. The interval without user activity that triggers a time-out from active use will impact the resulting EHR use times. Currently, this interval may vary from 5 seconds, to 90 seconds, to 30 minutes. Determining time-out intervals that best represent user activity is challenging and may vary by setting or task, and one could even argue that they should not be included at all in some calculations. For example, when a physician is working from home and multitasking with family obligations, all of the EHR time, both active and idle, potentially adversely impacts the physician and the quality of family time.

2. Definition of work outside of work. Some vendors and researchers have used clock times, such as 7 PM to 7 AM, to establish the window for work outside of work, but this does not account for variability in physician schedules. A physician who utilizes administration days or personal days off to complete documentation and inbox work would appear to have little WOW in this data model. In the ambulatory setting, an accurate assessment of WOW requires integration of the physician’s clinic schedule with the EHR log data.

3. The exclusion of “shoulder time” (ie, prep time) immediately before or after clinic. Some stakeholders have excluded the 30 minutes before and after scheduled patient time from WOW calculations. This minimizes the apparent WOW for those who work shorter clinic sessions. For example, a physician with 16 hours of patient contact time per week distributed in 2-8 hour days would have 2 hours of shoulder time subtracted from WOW, whereas a physician who distributed the same patient contact time across four 4-hour days would have 4 hours of shoulder time subtracted from WOW. To avoid these sources of distortion in the data, we define WOW precisely as that time on the EHR outside of scheduled patient hours.

4. The mapping of EHR actions (ie, time on a type of EHR screen) to work (ie, the total time required to complete a task). For example, the time to resolve an inbox task is not fully captured by the time on the inbox screen alone—it may also include reviewing the problem list, consulting the medication list, and assessing recent lab results. Our inbox measure captures the entire time on the task of inbox management as opposed to just the time spent on the inbox screen itself. The technical methodology for linking the time on multiple screens to capture time per task requires refinement.

5. Work outside of the EHR. Not every aspect of clinical work is performed via the EHR. The time associated with phone calls, messages, family meetings, paperwork, verbal communication with colleagues, and even direct face time with patients are components of a physician’s weekday that are not directly captured by EHR time stamp data.

VALIDITY

6. EHR use measure validation. Options include face validity via end-user testing vs validation against an external standard, such as direct observation time-motion analysis. Manual time-motion observations are labor-intensive but currently represent the gold standard for validation. It is our hope that these core measures will be validated across different types of institutions, clinical specialties, and EHR vendors.

NORMALIZATION

7. EHR measure normalization. Candidate denominators include number of patients, clinic sessions, workdays, or hours of scheduled patient time. We elected to normalize to 8 hours of patient scheduled time because time is a finite resource for physicians. We acknowledge that volume is a consideration as well; a physician who is scheduled lightly during a clinic session (ie, 1 patient per hour) may be able to perform more documentation, inbox, and other EHR work during their session than is a physician who is scheduled more densely during their session (ie, 4 patients per hour). Therefore, we encourage researchers and vendors to report average clinic volumes and, as feasible, other measures of workload, along with the time-based measures.

GENERALIZABILITY

9. Generalization to inpatient and procedural settings. Additional work is needed to adapt these measures to other settings. For example, when hospitalists and emergency room physicians work shift sets, WOW can be easily defined as any time outside of a scheduled shift and normalized to 8 hours of shift work. On the other hand, when physicians work in both ambulatory and hospital settings, it may be difficult to identify which WOW belongs to which setting.

10. Unintended negative consequences. Some physicians may not like the notion that “the company” is watching what they do at home or when they are not scheduled with the EHR log data. This could motivate them to shorten notes, cut corners, or find gaps in the amount of work required to collect and clean a larger dataset and the potential for a shorter period to reflect anomalies in patient care due to seasonal care requirements, vacations, time on teaching service, changes in staffing, etc.

CONCLUSION

The analysis of EHR log data is emerging as a tool to further understand the clinical environment and to optimize operational, technological, and policy decisions. We propose and define 7 core measures of EHR use and normalize the time measures to 8 hours of patient scheduled time. It is our intent that the measures provide insight and facilitate research regarding the efficiency of using EHRs in the practice environment, the effectiveness of teams, the impact of policies and regulations, and practice characteristics that contribute to physician distress or well-being. Improving the physician experience should, in turn, positively impact the patient experience as well. Furthermore, these standard measures will allow for better reproducibility and comparison of research studies. While we expect EHR use measures will need to adapt over time in response to advances in technology, changing clinician roles, and evolving regulatory policies, these measures provide a starting point to enhance consistency and reproducibility for analyses of current systems.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTIONS

Each author made substantial contributions to the conception or design of the work; was involved in drafting the work or revising it.
critically important for important intellectual content; gave final approval of the version to be published; and has agreed to be accountable for all aspects of the work.

CONFLICT OF INTEREST

SLB acknowledges grant funding from National Institutes of Health/ National Library of Medicine T15LM011271 and the Heed Ophthalmic Foundation. The remaining authors have no competing interests to declare.

REFERENCES

1. Thune J, Alexander L, Roberts P, Burr R, Enzi M. Where is HITECH's interest to declare. Health Aff (Millwood) 2011; 30 (3): 464–71.
2. Black AD, Car J, Pagliari C, et al. The impact of eHealth on the quality and safety of health care: a systematic overview. PLoS Med 2011; 8 (1): e1000387.
3. Kroth PJ, Morioka-Douglas N, Veres S, et al. Association of electronic health record design and use factors with clinician stress and burnout. JAMA Netw Open 2019; 2 (8): e199609.
4. Shanafeld TD, Dyrbye LN, Sinsky C, et al. Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction. Mayo Clin Proc 2016; 91 (7): 836–48.
5. Gardner RL, Cooper E, Haskell J, et al. Physician stress and burnout: the impact of health information technology. J Am Med Inform Assoc 2019; 26 (2): 106–14.
6. Sinsky C, Colligan L, Li L, et al. Allocation of physician time in ambulatory practice: A time and motion study in 4 specialties. Ann Intern Med 2016; 165 (11): 753–60.