**RESEARCH ARTICLE**

# Documentation dynamics: Note composition, burden, and physician efficiency

**Nate C. Apathy PhD** [1,2] | **Lisa Rotenstein MD, MBA** [3,4] |
**David W. Bates MD, MSc** [3,5] | **A. Jay Holmgren PhD, MHI** [6]

[1] National Center for Human Factors in Healthcare, MedStar Health Research Institute, Washington, District of Columbia, USA

[2] Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, Indiana, USA

[3] Harvard Medical School, Boston, Massachusetts, USA

[4] Population Health Brigham & Women's Hospital, Boston, Massachusetts, USA

[5] Division of General Internal Medicine, Brigham & Women's Hospital, Boston, Massachusetts, USA

[6] Center for Clinical Informatics and Improvement Research, University of California – San Francisco School of Medicine, San Francisco, California, USA

**Correspondence**
Nate C. Apathy, National Center for Human Factors in Healthcare, MedStar Health Research Institute, Washington, DC, USA.
Email: nate.apathy@medstar.net

**Present address**
David W. Bates, Department of Health Policy and Management, Harvard School of Public Health, Boston, MA, USA.

## Abstract

**Objective:** To analyze how physician clinical note length and composition relate to electronic health record (EHR)-based measures of burden and efficiency that have been tied to burnout.

**Data Sources and Study Setting:** Secondary EHR use metadata capturing physician-level measures from 203,728 US-based ambulatory physicians using the Epic Systems EHR between September 2020 and May 2021.

**Study Design:** In this cross-sectional study, we analyzed physician clinical note length and note composition (e.g., content from manual or templated text). Our primary outcomes were three time-based measures of EHR burden (time writing EHR notes, time in the EHR after-hours, and EHR time on unscheduled days), and one measure of efficiency (percent of visits closed in the same day). We used multivariate regression to estimate the relationship between our outcomes and note length and composition.

**Data Extraction:** Physician-week measures of EHR usage were extracted from Epic's Signal platform used for measuring provider EHR efficiency. We calculated physician-level averages for our measures of interest and assigned physicians to overall note length deciles and note composition deciles from six sources, including templated text, manual text, and copy/paste text.

**Principal Findings:** Physicians in the top decile of note length demonstrated greater burden and lower efficiency than the median physician, spending 39% more time in the EHR after hours ($p < 0.001$) and closing 5.6 percentage points fewer visits on the same day ($p < 0.001$). Copy/paste demonstrated a similar dose/response relationship, with top-decile copy/paste users closing 6.8 percentage points fewer visits on the same day ($p < 0.001$) and spending more time in the EHR after hours and on days off (both $p < 0.001$). Templated text (e.g., Epic's SmartTools) demonstrated a non-linear relationship with burden and efficiency, with very low and very high levels of use associated with increased EHR burden and decreased efficiency.

**Conclusions:** "Efficiency tools" like copy/paste and templated text meant to reduce documentation burden and increase provider efficiency may have limited efficacy.

**KEYWORDS**
documentation, electronic health records, health policy, physician burnout

**What is known on this topic**

- Electronic health record (EHR)-derived burden is a major driver of physician burnout, especially burden associated with clinical documentation.
- Time in the EHR after hours, EHR use on off days, and same day visit closure rates have all been correlated to physician burnout.
- Efforts to reduce burnout (e.g., the 2021 Evaluation and Management billing guideline changes) have focused in part on reducing documentation requirements and changing clinical notes.

**What this study adds**

- Our study offers the first and largest-scale national analysis of how physician note composition relates to measures of physician burden and efficiency that have been tied to burnout.
- Physicians that write the longest notes use the EHR more during off hours and on days off, and close fewer visits in the same day (i.e., experience more burden and less efficient).
- High use of the tools intended to improve documentation efficiency like templated text and copy/paste do not correlate with decreased burden or improved efficiency.

# 1 | INTRODUCTION

Following the passage of the HITECH Act of 2009,[1] acute care delivery organizations have rapidly adopted electronic health records (EHRs) in response to federal incentives.[2] While EHRs have promised significant improvements to care quality and reductions in costs, broad adoption of EHRs in ambulatory care has led to increased clerical burden and detrimental effects for the clinician experience.[3] Studies estimate that between 50% and 80% of physicians experience symptoms of burnout, and time spent working in the EHR is often cited as a driver of this phenomenon.[4–6] While the causes of burnout are numerous and complex, there is strong evidence that greater EHR burden has been associated with a higher risk of emotional exhaustion[7] and perceptions of clerical burden.[8,9] On average, physicians spend nearly a third of their EHR time on documentation and writing notes, with questionable clinical value given concerns over "note bloat" and burdensome administrative requirements contributing to the documentation work known as "desktop medicine."[10–12] The prevalence of this onerous clerical work—combined with its dubious clinical value—has led to broad policy changes, such as those championed by the American Medical Association (AMA) and Centers for Medicare and Medicaid Services (CMS) to common evaluation and management (E/M) billing codes for the explicit purpose of reducing documentation requirements for billing, to alleviate some of this clerical burden and improve physician efficiency by striking at one of the root causes of clinician burnout.[13–15] Reducing EHR documentation burden at the organizational level has been a focus of several health systems through a variety of interventions, such as outsourcing work to virtual scribes or implementing one-on-one training sessions with clinicians.[16–21]

Despite this focus on reducing documentation burden from both policy makers and health system leaders, little is known about the relationship between EHR documentation and physician efficiency and productivity. While the 2021 coding changes reduced documentation required to justify E/M billing codes, early evidence suggests it did not immediately result in a reduction in EHR note length, despite this policy's broad reach across payers and billing codes used in more than half of outpatient visits.[22] Furthermore, the 2021 E/M policy specifically targeted note length by reducing what a note must contain, but note length and note composition (i.e., how notes are written) may be related to physician burden and efficiency in distinct ways. For example, note *length* may impact efficiency in that certain care management workflows in primary care or inclusion of specific content by specialists may simply necessitate longer notes that lead to more EHR documentation time and lower physician efficiency while also worsening quality by making it harder to find key information. However, note *composition* may exert a different influence; for example, automated tools that assist in documentation may speed the process for specialists and improve these physicians' efficiency, but may be less effective for primary care physicians that see a more diverse array of clinical cases, rendering templates less universally applicable. Moreover, note composition style is likely upstream of note length, with long notes resulting from certain practices like extensive copy/paste or use of bloated text templates. It may also be the case that some physicians are more facile at using the EHR and make better use of automated tools demonstrate greater efficiency, independent of note length.[23] Additionally, some clinicians may be simply less aware or less well-versed in the use of these automated documentation tools, and document in less efficient ways. Affecting both, many EHRs lack good search tools to identify relevant note templates. Finally, these relationships may be non-linear; it may be that some amount of templated text improves efficiency, but overreliance on automated tools reduces flexibility and results in bloated notes that slow down clinicians and reduce overall efficiency. Lacking empirical evidence of these relationships, however, policy makers and health system leaders are largely flying blind with respect to the highest-impact levers to pull to reduce documentation burden and improve physician efficiency.

Several studies have analyzed dimensions of EHR use and burnout,[24] and have reliably shown that more time in the EHR—especially time spent in the domains of inbox management and after-hours time—is positively associated with burnout.[6,7,25,26] Burnt out physicians are also less efficient in terms of same-day visit closure rates.[25] Despite this evidence, linking note length and/or composition directly to burnout has been difficult, and thus far the findings are mixed.[6,23] In one study, note length was not correlated with burnout, while copy/paste use counter-intuitively appeared to be protective against burnout.[6] While most of the evidence to-date has focused on time spent in the EHR as an antecedent to burnout, it is important to recognize that total physician time in the EHR is difficult to directly intervene on. Therefore, an understanding of the mechanisms contributing to extended time in the EHR for documentation and physician efficiency is critical for identifying intervenable aspects of physician EHR use that have the potential to reduce burnout.

To address this gap, we sought to characterize the relationship between both note length and note composition and aspects of physician EHR burden correlated with burnout in ambulatory care. Using a unique national dataset, our study addresses three research questions. First, how does overall note length relate to physician EHR burden and efficiency? Second, how does note composition (e.g., use of manual entry, copy/paste, and templated text) relates to physician burden and efficiency? And finally, do these relationships vary across specialties? Quantifying these relationships and any specialty variation may inform cross-specialty learning and help to disseminate documentation best practices, as well as highlight specialties that may be under-served by existing EHR tools. More broadly, a greater understanding of the mechanisms underlying the factors correlated with physician burnout can help inform health system and EHR vendor investment in documentation features and tool usage to directly reduce burden, with second-order implications for physician burnout. Finally, these questions can shape CMS and other payers' policies aiming to reduce physician burnout via changes to how physicians document in the EHR.

## 2 | METHODS

### 2.1 | Data source

Data for this study were derived from a national database of all outpatient physicians who used Epic Systems' ambulatory EHR from September 2020 through May 2021. All data came from weekly physician measures available through the Epic signal data warehouse[12,27] and were aggregated to physician-level averages for measures of EHR time, note length, note composition, and visit closure rates for 203,728 physicians across 392 health systems. Measures represent only time spent in the EHR for ambulatory care; inpatient EHR time is not observed. This effectively represents Epic's entire US physician user-base during this period. All physicians and organizations were de-identified prior to receipt of the data, and the [University of Pennsylvania] Institutional Review Board deemed this study exempt as non-human subjects research.

### 2.2 | Measures: Note length and note composition

To measure note length, our primary independent variable was average note length per note, measured in characters per note. This measure is calculated by dividing total characters authored in notes by the physician by the total number of notes the physician wrote. This methodology is consistent with other studies using EHR audit log data derived from Epic's Signal data extraction tool.[12,13] We also computed physician-level deciles of this measure for use in regression models, described below.

To measure note composition, we identified the average number of characters per note attributable to each of six documentation methods identified in Signal: manual note entry; SmartTools (templated text also known as "dot phrases"); NoteWriter (templated text generated from a point-and-click graphical user interface); copy and paste; transcription; and voice recognition. For example, a 2000-character note might have 1000 characters from SmartTools (e.g., templated text or semi-structured forms), 500 characters from manual entry (e.g., the physician entering information by simply manually typing), and 500 characters from copy and paste (e.g., content copied from other notes or result reports). We binned physicians into deciles according to the full-sample distribution of the amount of per note content from each source for our regression analysis.

### 2.3 | Measures: Physician burden and efficiency

#### 2.3.1 | EHR time and use on days off

We examined three outcome variables capturing measures that previous literature has found to be correlated with physician burnout: time in notes per visit, time after hours per visit, and unscheduled days with EHR use.

To measure time in the EHR, Epic calculates a measure of "active time" in the EHR for each physician-week as the sum of time during which the physician was actively engaged with the EHR system. Epic measures active time across several domains of the EHR, including in note-writing activities, clinical review of patient data, and inbox management. Previous studies have explored in more detail the precise methodology used to calculate this measure and how it compares to other EHR vendors' methods.[12,27,28] Generally speaking, Epic's measure can be interpreted as a lower-bound estimate of time spent in the EHR. To account for differences in physician scheduling, aggregate time in each domain (notes and after hours) was divided by total visits, to arrive at measures of "active EHR time [*in notes/after-hours*] per visit." Off days were measured as the average number of days per week during which the physician had no scheduled visits but logged time in the EHR.
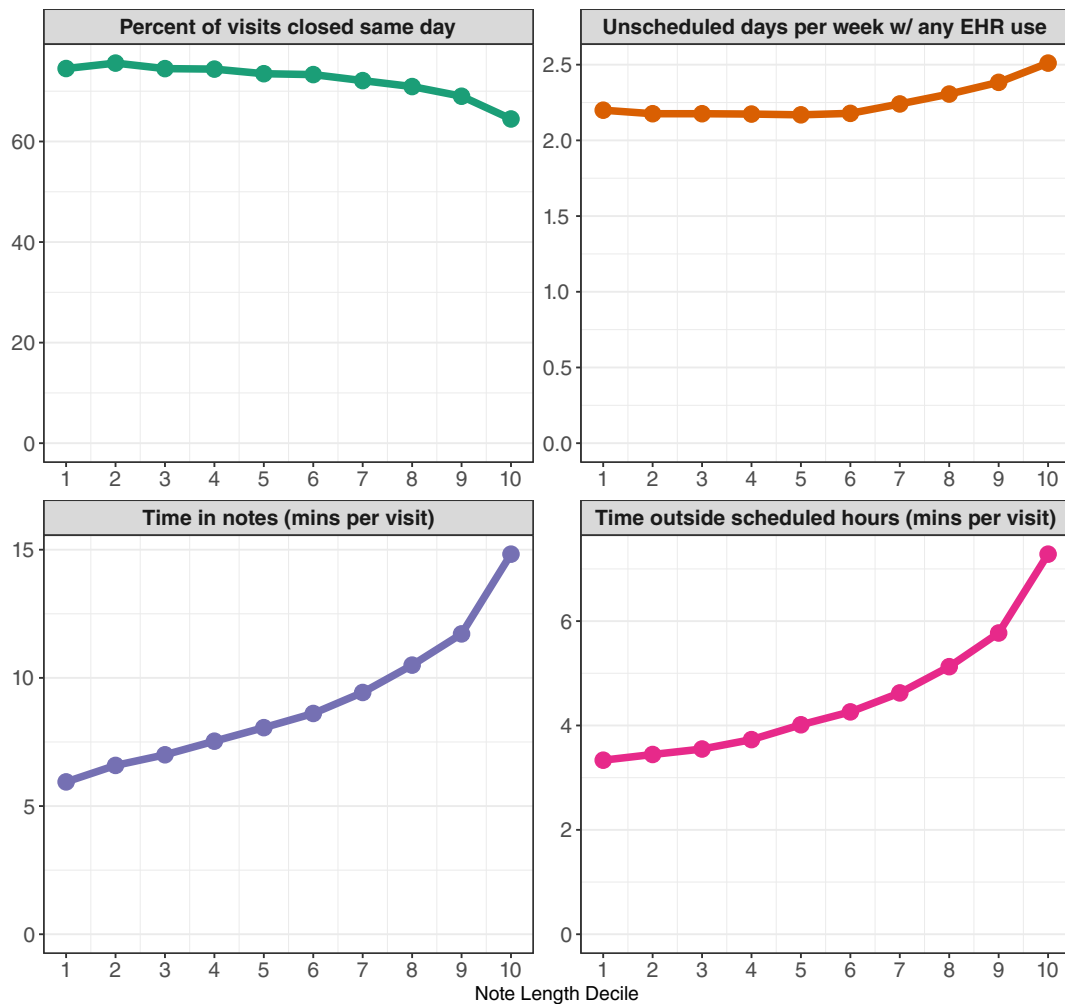
**FIGURE 1** Average physician burden and efficiency measures by overall note length decile. *Source*: Authors' analysis of outpatient EHR usage data for 203,728 US-based physicians using the Epic EHR between September 2020 and May 2021. [Color figure can be viewed at wileyonlinelibrary.com]

## 2.3.2 | Physician efficiency

Our fourth outcome measure was visit closure efficiency, measured as the percentage of a physician's visits that were closed on the same day as the visit took place. We chose this measure following the logic that a lower percentage of visits closed on the same day indicates either physician inefficiency in EHR use, overly burdensome EHR design that prevents efficient use, or some combination of these two forces. As with all other measures, this was averaged at the physician level over the study period for analyses.

## 2.3.3 | Physician and organization covariates

To facilitate comparisons across specialties, physicians included in our analysis were grouped into primary care, medical specialties, or surgical specialties consistent with previous work examining specialty variation in EHR time.[29] We excluded several specialties such as anesthesiology that are not representative of general outpatient practice.

We used several measures to account for physician case complexity, which may influence both note length and physician burden. These included average patient age, average number of weekly visits, and five measures capturing the percent of weekly E/M visits billed to new and established E/M levels one through five. Consistent with prior work,[12,29] and in order to preserve organizational anonymity, we only observe structure (ambulatory-only; hospital and clinic facilities; other) and US census region.

## 2.3.4 | Analysis

We first computed descriptive statistics across all measures for the sample overall and stratified across the three specialty groups. We then computed averages for all four outcome measures at each note length decile and visualized these relationships.

For our first research question to analyze the relationship between our outcomes of physician burden and overall note length, we used multivariate ordinary least-squares regression models regressing our four outcomes on a categorical variable capturing physician
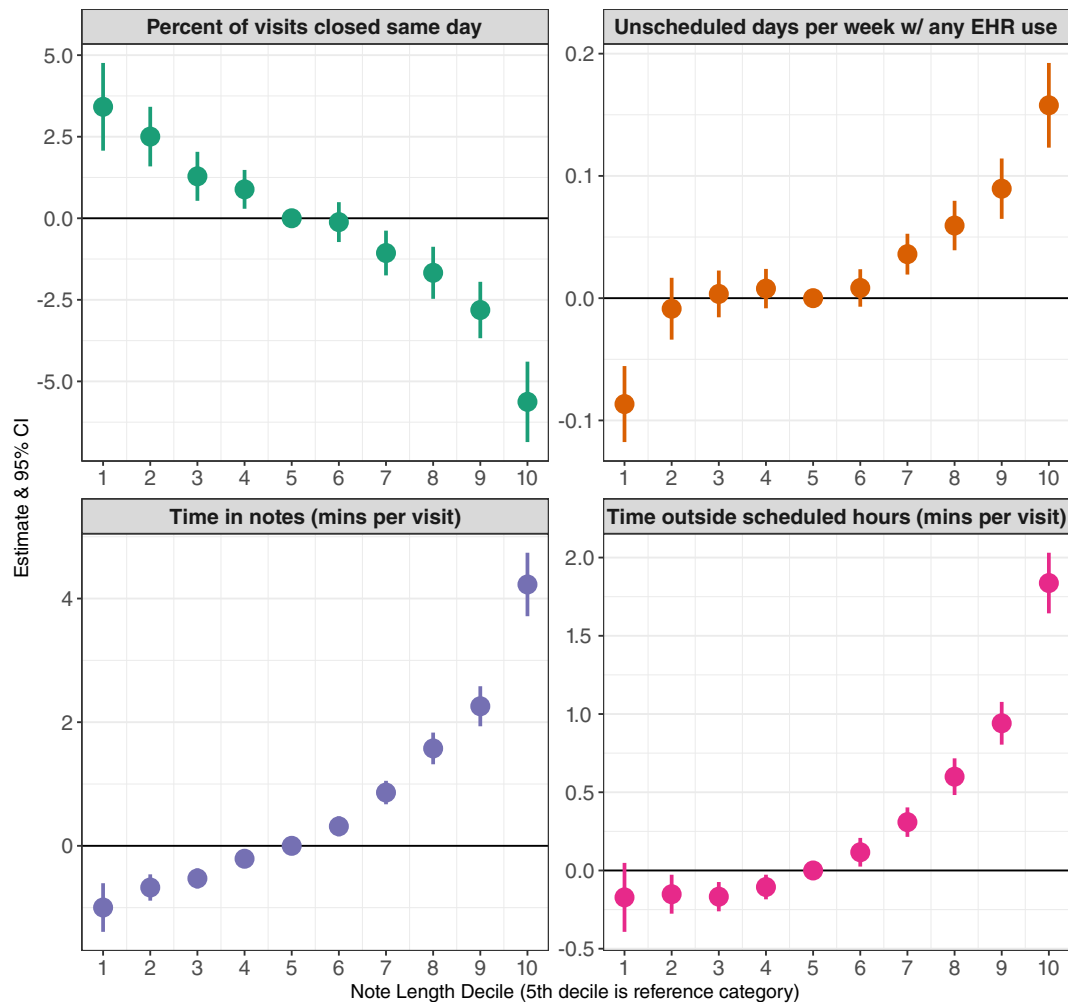
**FIGURE 2** Relationship between overall note length and physician burden and efficiency measures. Each row displays findings from a single multivariable ordinary least-squares regression model for each outcome adjusting for physician case load and case mix complexity, and organizational factors. The fifth decile is used as the reference category to obtain estimates for physicians in each decile of note length. Each of the four outcome panels represent some aspect of physician burden. Full regression results for all models can be found in Appendix Table 1. [Color figure can be viewed at wileyonlinelibrary.com]

deciles of overall note length, with the fifth decile as the reference category. For our second research question to analyze the relationship between our outcomes and note composition, we used multivariate ordinary least-squares regression models regressing our four outcomes on six categorical variables, each capturing physician-level deciles of note characters from each of the six sources of note text: manual note entry; copy and paste; SmartTools; Note-Writer; transcription; and voice recognition. For these independent variables, the fifth decile was the reference group. For our final research question examining differences across specialties, we re-ran all regression models stratified by physician specialty group, for a total of 24 stratified models. All models included covariates to adjust for average patient age, visit volume, E/M visit mix, organizational type, and US census region for the organization.

All models used two-way clustered errors at the physician and organization levels, and we used an $\alpha$ value of 0.05 to gauge statistical significance. For all regression results, we present coefficient plots showing the estimates and 95% confidence intervals for our

independent variable(s) of interest in each model. All analyses were conducted in R version 3.6.0 using RStudio.[30,31] The *fixest* package was our primary analytical package.[32] Data management was done using the *tidyverse* suite of tools and the *lubridate* package.[33,34]

## 3 | RESULTS

### 3.1 | Overall note length and burden

For all four outcomes of interest (EHR time in notes per visit, EHR time after hours per visit, time in the EHR on unscheduled days, and percent of visits closed on the same day), physicians in the top deciles of overall note length demonstrated more burden (Figure 1). Physicians in the top decile of overall note length spent more time in notes per visit than physicians in the median (fifth) decile of note length (14.8 vs. 8.1 min per visit). Similarly, top decile physicians spent more time in the EHR per visit outside scheduled hours (7.3 vs. 4.0 min) and had more

**TABLE 1** Descriptive statistics of sample, overall, and stratified by specialty group

| | Overall | Primary care | Medical specialties | Surgical specialties |
|---|---|---|---|---|
| # of physicians | 203,728 | 66,968 | 80,267 | 56,493 |
| Physician burden and efficiency measures | | | | |
| % of visits closed in the same day | 72.1 (29.7) | 76.5 (26.5) | 66.8 (32.2) | 74.5 (28.5) |
| Time in notes per visit | 9.1 (10.5) | 9.3 (9.4) | 11.6 (11.7) | 5.5 (8.8) |
| Time outside scheduled hours per visit | 4.6 (5.5) | 4.8 (5.7) | 5.7 (6.1) | 2.6 (3.5) |
| Unscheduled days with any EHR use per week | 2.3 (1.0) | 2.0 (0.8) | 2.3 (1.0) | 2.5 (1.0) |
| Documentation length (characters per note) | 4835.5 (3045.3) | 4199.8 (2208.1) | 6128.8 (3574.3) | 3751.5 (2341.5) |
| Note composition (characters per note) | | | | |
| Copy/paste | 749.5 (1580.8) | 245.8 (747.9) | 1433.7 (2171.2) | 374.6 (737.7) |
| SmartTools | 1820.0 (1527.0) | 1889.6 (1413.4) | 2070.3 (1674.8) | 1381.9 (1329.9) |
| Manual text | 363.0 (419.8) | 330.3 (347.8) | 472.9 (514.7) | 245.6 (288.7) |
| NoteWriter | 208.8 (442.7) | 350.6 (533.9) | 158.5 (419.4) | 112.3 (286.5) |
| Transcription | 37.7 (335.1) | 11.0 (188.0) | 54.0 (398.1) | 46.2 (369.4) |
| Voice recognition/dictation | 227.9 (455.8) | 168.9 (371.0) | 268.6 (520.7) | 239.8 (440.8) |
| Average patient age | 48.6 (19.0) | 44.1 (21.3) | 51.4 (18.8) | 50.1 (15.2) |
| Total weekly visits | 35.2 (26.3) | 44.8 (27.4) | 27.3 (23.1) | 34.9 (25.5) |
| % of E/M visits billed to Level 1 | 0.7 (4.5) | 0.4 (2.7) | 0.7 (4.5) | 1.2 (6.0) |
| % of E/M visits billed to Level 2 | 3.4 (8.0) | 2.7 (6.3) | 2.2 (6.6) | 5.9 (10.5) |
| % of E/M visits billed to Level 3 | 20.7 (18.8) | 24.1 (17.3) | 15.4 (17.9) | 24.1 (19.8) |
| % of E/M visits billed to Level 4 | 27.7 (22.7) | 29.3 (20.3) | 34.3 (24.7) | 16.6 (17.8) |
| % of E/M visits billed to Level 5 | 7.7 (14.8) | 3.6 (9.1) | 13.8 (18.9) | 3.9 (10.0) |
| US census region (n, %) | | | | |
| Midwest | 57,965 (28) | 19,182 (29) | 22,399 (28) | 16,384 (29) |
| Northeast | 45,318 (22) | 13,386 (20) | 20,210 (25) | 11,722 (21) |
| South | 50,136 (25) | 15,707 (23) | 19,785 (25) | 14,644 (26) |
| West | 50,309 (25) | 18,693 (28) | 17,873 (22) | 13,743 (24) |
| Organizational type (n, %) | | | | |
| Ambulatory-only | 14,492 (7) | 7043 (11) | 4067 (5) | 3382 (6) |
| Hospital and clinic facilities | 179,459 (88) | 56,083 (84) | 72,893 (91) | 50,483 (89) |
| Other | 9777 (5) | 3842 (6) | 3307 (4) | 2628 (5) |

*Note*: All values are mean (SD) unless noted otherwise.
Abbreviations: EHR, electronic health record; E/M, evaluation and management.

unscheduled days per week with any EHR use (2.5 vs. 2.2 days). These physicians also demonstrated lower efficiency, closing only 64.5% of visits on the same day, in comparison to 73.5% of visits closed in the same day by median note length physicians.

These dose–response relationships held in our regression analyses adjusting for physician visit volume, case complexity, and organizational factors, with the top decile of note length reliably associated with greater burden measures (Figure 2). Compared to median note length physicians, being in the top decile of overall note length was associated with 4.2 more minutes per visit spent on notes ($B = 4.22$; 95% CI: 3.71–4.74; $p < 0.001$), representing 46% more time in notes than the sample average of 9.1 min per visit (Table 1). These physicians also spent 1.8 more minutes per visit in the EHR outside scheduled hours ($B = 1.84$; 95% CI: 1.64–2.03; $p < 0.001$), 39% longer than the average physician. Top decile physicians also had 0.16 more

unscheduled days per week with any EHR use in comparison to median physicians ($B = 0.16$; 95% CI: 0.12–0.19; $p < 0.001$) and closed 5.6 percentage points fewer visits on the same day as the visit ($B = -5.63$; 95% CI: $-6.86$ to $-4.40$; $p < 0.001$). These represent increases of 6.9% and 7.8% compared to the average physician, respectively. Full results of our note length decile regressions for all four outcomes can be found in Appendix Table 1.

## 3.2 | Note composition and burden

### 3.2.1 | Time in the EHR

For our analysis using note composition tools as our independent variables of interest, we focus on reporting results on the relationships
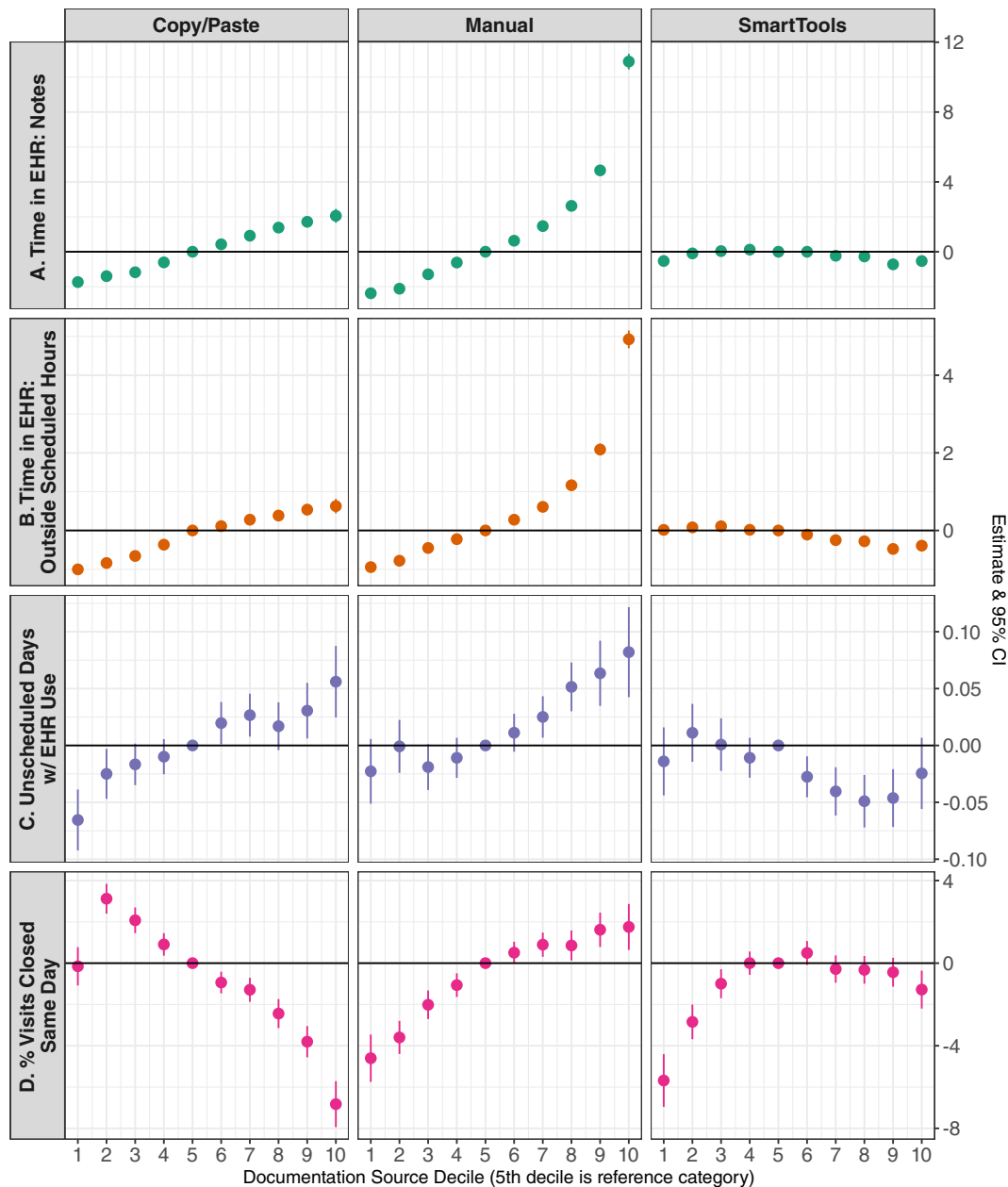
**FIGURE 3** Relationship between note composition tool use and physician burden and efficiency. Each row displays findings from a single multivariable ordinary least-squares regression model adjusting for additional deciles of note source (transcription, voice recognition, and NoteWriter), physician case load (average visits per week) and case mix complexity, and organizational factors. For each note text source, the fifth decile is used as the reference category to obtain estimates for physicians in each decile of note content from each of the three featured text sources. Each of the four outcomes on the left side represent some aspect of physician burden. Units for panels (A, B) are minutes per visit; unit for panel (C) is days per week; unit for Panel (D) is percentage points. Full regression results for all models can be found in Appendix Table 2. [Color figure can be viewed at wileyonlinelibrary.com]

between our four outcomes and (1) copy/paste text, (2) manual text, and (3) text from SmartTools, as these three sources account for more than two-thirds of note text, on average. Physician time in notes increased with increased note text from both copy/paste text and manual text but was not higher for physicians with the most note text from SmartTools (Figure 3). Most dramatically, physicians in the top decile of manual text per note spent 10.9 more minutes in notes per visit than physicians with median use of manual note text ($B = 10.89$; 95% CI: 10.45–11.34; $p < 0.001$; a 147.0% relative increase).

Conversely, physicians in the top decile of SmartTools text spent 0.53 fewer minutes in notes than median SmartTools users ($B = -0.5329$; 95% CI: $-0.8421$ to $-0.2236$; $p < 0.001$; a 5.6% relative decrease).

Similar patterns were present for time in the EHR outside of scheduled hours, although the magnitudes of these estimates were smaller. Top-decile manual text physicians spent, on average, 4.9 more minutes in the EHR outside of scheduled hours per visit than median manual text users ($B = 4.92$; 95% CI: 4.70–5.15; $p < 0.001$). As with time in notes, physicians in the top decile of SmartTools use illustrated

**TABLE 2** Relationship between note composition tool use and physician burden and efficiency, selected results from specialty-stratified regression models

| | Unscheduled days with any EHR use | | | |
| --- | --- | --- | --- | --- |
| | Overall<br>B [95% CI] | Primary care<br>B [95% CI] | Medical specialties<br>B [95% CI] | Surgical specialties<br>B [95% CI] |
| Copy/paste text | | | | |
| Decile 1 | **−0.07 [−0.09, −0.04]**\*** | −0.03 [−0.06, 0] | **−0.07 [−0.13, −0.02]*** | **−0.12 [−0.16, −0.07]**\*** |
| Decile 5 (ref) | | | | |
| Decile 10 | **0.06 [0.02, 0.09]**\*** | **0.14 [0.07, 0.22]**\*** | **0.07 [0.03, 0.11]**\*** | −0.04 [−0.1, 0.02] |
| Manual text | | | | |
| Decile 1 | −0.02 [−0.05, 0.01] | **−0.06 [−0.1, −0.02]**** | −0.04 [−0.09, 0] | −0.02 [−0.06, 0.02] |
| Decile 5 (ref) | | | | |
| Decile 10 | **0.08 [0.04, 0.12]**\*** | **0.19 [0.14, 0.25]**\*** | −0.03 [−0.09, 0.03] | **0.15 [0.08, 0.23]**\*** |
| SmartTools text | | | | |
| Decile 1 | −0.01 [−0.04, 0.02] | −0.01 [−0.05, 0.03] | −0.04 [−0.08, 0] | 0.04 [−0.01, 0.08] |
| Decile 5 (ref) | | | | |
| Decile 10 | −0.02 [−0.06, 0.01] | −0.03 [−0.07, 0.01] | −0.03 [−0.07, 0.02] | −0.02 [−0.08, 0.03] |
| | % of visits closed same day | | | |
| | Overall | Primary care | Medical specialties | Surgical specialties |
| Copy/paste text | | | | |
| Decile 1 | −0.15 [−1.08, 0.78] | **3.03 [1.84, 4.21]**\*** | **−4.49 [−6.08, −2.91]**\*** | **−1.52 [−2.94, −0.11]*** |
| Decile 5 (ref) | | | | |
| Decile 10 | **−6.83 [−7.94, −5.71]**\*** | **−7.1 [−9.28, −4.92]**\*** | **−6.36 [−7.65, −5.08]**\*** | **−9.31 [−11.37, −7.26]**\*** |
| Manual text | | | | |
| Decile 1 | **−4.6 [−5.75, −3.46]**\*** | **−2.59 [−3.95, −1.24]**\*** | **−2.55 [−4.29, −0.81]**** | **−4.36 [−5.99, −2.73]**\*** |
| Decile 5 (ref) | | | | |
| Decile 10 | **1.75 [0.64, 2.87]**** | **−1.7 [−3.15, −0.24]*** | **6.23 [4.66, 7.79]**\*** | −1.33 [−2.98, 0.32] |
| SmartTools text | | | | |
| Decile 1 | **−5.68 [−6.96, −4.41]**\*** | **−7.31 [−10.38, −4.23]**\*** | **−3.08 [−4.89, −1.27]**\*** | **−5.61 [−7.38, −3.84]**\*** |
| Decile 5 (ref) | | | | |
| Decile 10 | **−1.28 [−2.2, −0.36]**** | −0.85 [−2.23, 0.52] | −0.93 [−2.3, 0.44] | **−2.43 [−4.21, −0.65]**** |

*Note*: Highlighted estimates reflect specialty deviations from overall estimates in either a normatively "better" (blue) or "worse" (red) direction. For example, primary care providers in the top decile of manual text demonstrate the reverse relationship from the overall average in the share of visits closed in the same day (1.7 pp fewer for PCPs vs. 1.8 pp more overall), so that estimate is highlighted in red.
*p < 0.05; **p < 0.01; ***p < 0.001.

slight time savings (B = −0.39 min per visit; 95% CI: −0.531 to −0.252; p < 0.001; an 8.3% relative savings), and high use of copy/paste text was associated with greater time spent in the EHR outside scheduled hours. Physicians using the most copy/paste text spent on average 14.4% more time in the EHR outside scheduled hours than median copy/paste users (B = 0.627 min per visit; 95% CI: 0.442–0.813; p < 0.001).

Finally, these patterns also held for the relationship between note text sources and the number of unscheduled days per week during which a physician had any EHR use. Relative to the median copy/paste and manual text physicians, physicians in the top deciles had 0.06 and 0.08 more unscheduled days with EHR use, respectively (B_copy/paste = 0.056; 95% CI: 0.025–0.088; p < 0.001 and B_manual = 0.082; 95% CI: 0.043–0.122; p < 0.001).

## 3.2.2 | Physician efficiency

High use of copy/paste text was associated with lower efficiency. Physicians in the top decile of copy/paste closed, on average, 6.8 percentage points fewer visits on the same day compared to median copy/paste physicians (B = −6.83; 95% CI: −7.94 to −5.71; p < 0.001; a 9.2% relative decrease). Conversely, physicians in the top decile of manual text use closed 1.8 percentage points *more* visits on the same day (B = 1.75; 95% CI: 0.64–2.87; p < 0.001; a 2.5% relative increase). SmartTools use did not demonstrate a strong relationship with physician efficiency. Regression estimates for deciles four through nine of SmartTools use were not statistically different from zero; however, physicians at the extremes of SmartTools use (i.e., in the bottom decile or top decile)

had lower efficiency, with bottom-decile SmartTools users closing on average 5.7 percentage points fewer visits than the median SmartTools user ($B = -5.68$; 95% CI: $-6.96$ to $-4.41$; $p < 0.001$), and those with very *high* use of SmartTools (in the top decile) closing 1.3 percentage points fewer visits ($B = -1.28$; 95% CI: $-2.20$ to $-0.036$; $p = 0.006$; a 1.7% relative decrease). Full results of our note composition decile regressions for all four outcomes can be found in Appendix Table 2.

## 3.3 | Variation across physician specialty groups

In stratified regressions, we found very little variation across physician groups in the relationships described above. Specialties did not vary in the relationships between overall note length and note composition and our two outcomes of per visit time in notes and time outside of scheduled hours. However, for the outcome of unscheduled days with any EHR use, some specialty differences emerged. Specifically, the relationship between high rates of copy/paste and manual text and more unscheduled days per week with EHR use is largely driven by primary care physicians. In stratified models, the relationship between this outcome and the top two deciles of copy/paste and manual text was roughly double the magnitude for primary care physicians, compared to the overall physician sample (Table 2). For example, we noted above that in the full sample, relative to the median manual text physicians, physicians in the top decile had 0.08 more unscheduled days with EHR use; for primary care physicians this estimate was 0.19 more unscheduled days with EHR use ($B = 0.19$; 95% CI: 0.14–0.25; $p < 0.001$).

There were also two notable specialty differences in our physician efficiency results. First, relative to the median physician, the lowest decile of copy/paste use had a positive relationship with physician efficiency among primary care physicians ($B = 3.03$; 95% CI: 1.84–4.21; $p < 0.001$) but a negative relationship among medical specialists ($B = -4.49$; 95% CI: $-6.08$ to $-2.91$; $p < 0.001$). In the overall sample, this relationship was not statistically different from zero. Finally, medical specialists in the top decile of manual text use appear to be driving the findings in the overall sample, with these physicians closing on average 6.2 percentage points more visits on the same day than median medical specialists ($B = 6.23$; 95% CI: 4.66–7.79; $p < 0.001$). To contextualize this result, it is important to note that medical specialists had on average the fewest visits per week (27.3 visits, Table 1), the longest notes (6129 characters per note), and the most manual text (473 characters per note).

## 4 | DISCUSSION

Physicians who write longer notes tend to exhibit more EHR-related burden, in the form of more time in the EHR after hours, on off days, and fewer visits closed on the day of the visit. These three measures are all associated with physician burnout,[6,7,24,26] suggesting that note length and note bloat play a key mechanistic role in the burnout crisis among US physicians. Note bloat has been attributed in part to billing

requirements and medico-legal pressures;[13] however, recent evidence illustrates that note length was unresponsive to relaxed documentation requirements in 2021[22] and is not correlated with aggregate measures of physicians' malpractice environment.[35] These findings call into question the salience of these oft-cited factors and suggest exploration of other socio-technical factors that may underpin bloated notes and frustrate efforts to alleviate burden. For example, physicians often lament the low usability of EHRs,[17,36,37] which creates burdensome information retrieval processes.[38,39] While EHRs are highly configurable by organizations[40] and clinicians can personalize their views to an extent, the clinical note offers a more-or-less infinitely customizable space for physicians to place any information important for their practice in an easily retrievable location. This creates duplicative information, fails to leverage structured data, and may lead to bloated notes that our findings suggest increase physician burden rather than decrease it. As a result, efforts to improve EHR information retrieval and align clinical information systems with clinician cognition and intuition[41,42] may help to reduce the dependence on notes as a "catch-all," thus reducing note bloat and alleviating at least some aspect of burden.

Our analysis of note composition suggests that some tools offer gains to efficiency and reduced burden, but some of these relationships are at odds with one another and demonstrate trade-offs. Physicians using low amounts of copy/paste and manual text spend less time in the EHR, both in notes and outside clinic hours—this may be in part due to these physicians writing shorter, less burdensome notes overall. However, these two note composition tools demonstrate opposing relationships with efficiency: less use of copy/paste is correlated with greater rates of same day visit closure while the opposite is true of manual text. Unfortunately, greater use of manual text is associated with both higher visit closure rates *and* more EHR time burden, implying that higher visit closure rates may come at a steep time cost and contribute to after-hours and off-day EHR use. Heavy reliance on copy/paste may reduce EHR documentation time during scheduled hours, but create a "pile-up" of work that reduces same-day visit closure efficiency. This mechanism is given support by our findings that higher levels of copy/paste are associated with increased time in the EHR outside of scheduled hours and EHR use during unscheduled days. Furthermore, stratified analyses show that this time burden is largely borne by primary care physicians, who may realize visit closure *inefficiencies* at high rates of manual text use in comparison to medical specialists.

Templated text demonstrated a nonlinear relationship with all four outcome measures, indicating that for efficiency in particular, use of SmartTools or documentation macros does not improve efficiency beyond the median level and may even hinder efficiency at very high levels of use. Moreover, our estimates for the association between SmartTools and outcomes were considerably smaller than for copy/paste or manual text, illustrating that both EHR time burden and efficiency are less closely tethered to templated text use relative to copy/paste and manual text use. Importantly, contrary to their branding as "efficiency tools," high use of copy/paste and templated text may exacerbate documentation-based burden. Templated text especially confers only small benefits with rapidly diminishing returns.

## 4.1 | Implications for policy and practice

Describing the relationships between different clinical documentation tools and physician EHR burden and efficiency help develop hypotheses regarding the mechanisms underpinning antecedents to burnout, namely administrative burden and after-hours EHR use. In comparison to these antecedents and burnout itself, note composition tools are relatively easy to intervene on via organizational efforts to reduce physician documentation burden. For example, while after-hours EHR use may be an important endpoint, our findings suggest that identifying physicians with high copy/paste use and tailoring documentation optimization efforts to that group may be effective in reducing after-hours EHR time. Additionally, our findings suggest initial gains but limited marginal returns to increased use of templated text, which may help organizations prioritize efforts that focus on driving adoption of SmartTools among non-users over efforts that seek to enhance SmartTools use among physicians that already use these tools.

Our results support policy efforts by CMS and other payers to limit documentation requirements to the information needed for clinical decision making and inter-provider communication, given their clear relationships with EHR burden and efficiency. The 2021 E/M guideline changes implemented by CMS sought directly to reduce administrative burden by reducing documentation requirements and shortening bloated clinical notes, however this policy stopped short of illustrating precisely *how* physicians might shorten notes while preserving guideline concordance. This ambiguity may have played a role in the tepid effect of the change on overall documentation length.[22] While it may be possible that the E/M guideline changes eventually result in reducing documentation time, this slow diffusion further illustrates the "stickiness" of existing documentation practices and habits. In future efforts, CMS and other payers may need to provide more prescriptive guidance on how clinical documentation can be shortened in order to overcome entrenched habits and workflows, as physicians are likely to continue to over-document in the presence of ambiguity.[43] Collaborating with EHR vendors and professional societies to develop and validate templates that can be easily implemented by health systems and quickly adopted by physicians might help to accelerate reductions in clinical note bloat and begin to ease this source of physician burden, a function federal policy makers could incentivize through levers such as the ONC EHR Reporting Program.

Going forward, policy makers should also consider prioritizing research efforts to identify additional salient pain points of EHR usability that drive burden and burnout, to expand the levers available to organizations seeking to reduce physician burden beyond those focused on documentation, which constitutes only a fraction of clinical practice.[44–46] For example, prior authorization is widely cited as a highly administratively onerous process, and all major EHRs have tools to make prior authorization more streamlined. Yet little is known about these systems, the degree to which they do in fact streamline the process, the usability pain points, or how they can be best designed to maximize efficiency. Future research should leverage randomized trials within organizations as well as natural experiments and evaluations of EHR burden reduction programs to more precisely identify the proximate causes of the note bloat that leads to physician EHR burden and ultimately burnout.[47,48]

## 4.2 | Limitations

Our study has several limitations. First, given the observational and cross-sectional nature of our study design, we are unable to comment on any causal relationship between note length, composition, and our outcomes of interest. However, we observe clear dose–response relationships that can help to inform organizational efforts to reduce physician burden. While previous studies have suggested overall note length has remained relatively consistent throughout the study period,[22,49] future research should evaluate individual note composition patterns longitudinally with a focus on sources of quasi-experimental variation to better identify the impact of changes in physician documentation practices on efficiency. Second, we do not observe note content and thus cannot comment on aspects of note quality, which may be related to use of copy/paste, SmartTools, and manual text. Third, our measure of physician burden and efficiency (share of visits closed in the same day) is a proxy, and may contain variation related to physician preferences (e.g., a preference to close out visits on the next day after follow-up calls can be made). Nevertheless, in aggregate and on average, this is a useful and normatively interpretable measure of physician administrative burden and efficiency, as all else equal, physicians would rather close out more visits in the same day. Fourth, given that our data is de-identified, we are unable to control for any organizational efforts to streamline documentation, reduce after-hours EHR use, or alleviate administrative burden. We are also unable to observe the modality (in-person vs. virtual telemedicine) of encounters, and there may be important differences in documentation patterns and relationships with efficiency between in-person care and telemedicine visits that should be explored in future research. Fifth, our data is drawn from ambulatory physicians using a single EHR vendor, Epic. However, Epic has the largest ambulatory care market share,[50] and is broadly implemented across a range of practice settings, including safety net providers.[51] Additionally, our sample includes conservatively more than one-tenth of the entire US physician workforce, which to our knowledge constitutes the largest single study of physician note composition to-date.

## 5 | CONCLUSION

We find a clear, linear relationship between physician note length and measures of EHR burden and efficiency, illustrating that physicians with the longest clinical notes reliably experience greater burden. However, not all note text is created equal with respect to burden and efficiency. Increasing use of copy/paste text reliably results in greater burden, but manual text and templated text demonstrate inverted and nonlinear relationships with efficiency, respectively. This evidence can help inform organizational efforts to reduce burden

via intervenable mechanisms like physician note composition styles, but future research will be required to more precisely identify the salient and causal features of EHR design that drive note bloat, burden, and physician burnout. In turn, those features can be redesigned to enable more effective clinical documentation that can improve quality and reduce costs.

## ACKNOWLEDGMENTS

## ORCID

*Nate C. Apathy* https://orcid.org/0000-0002-7775-8544

## REFERENCES

1. Blumenthal D. Launching HITECH. *N Engl J Med*. 2010;362(5): 382-385.
2. Apathy NC, Holmgren AJ, Adler-Milstein J. A decade post-HITECH: critical access hospitals have electronic health records but struggle to keep up with other advanced functions. *J Am Med Inform Assoc*. 2021; 28:1947-1954. doi:10.1093/jamia/ocab102
3. Lin SC, Jha AK, Adler-Milstein J. Electronic health records associated with lower hospital mortality after systems have time to mature. *Health Aff*. 2018;37(7):1128-1135.
4. West CP, Dyrbye LN, Erwin PJ, Shanafelt TD. Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. *Lancet*. 2016;388(10057):2272-2281.
5. Rotenstein LS, Torre M, Ramos MA, et al. Prevalence of burnout among physicians: a systematic review. *JAMA*. 2018;320(11):1131-1150.
6. Hilliard RW, Haskell J, Gardner RL. Are specific elements of electronic health record use associated with clinician burnout more than others? *J Am Med Inform Assoc*. 2020;27:1401-1410. doi:10.1093/jamia/ocaa092
7. Adler-Milstein J, Zhao W, Willard-Grace R, Knox M, Grumbach K. Electronic health records and burnout: time spent on the electronic health record after hours and message volume associated with exhaustion but not with cynicism among primary care clinicians. *J Am Med Inform Assoc*. 2020;27(4):531-538.
8. Shanafelt 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):836-848.
9. 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):106-114.
10. Tai-Seale M, Olson CW, Li J, et al. Electronic health record logs indicate that physicians split time evenly between seeing patients and desktop medicine. *Health Aff*. 2017;36(4):655-662.
11. Overhage JM, McCallie D Jr. Physician time spent using the electronic health record during outpatient encounters: a descriptive study. *Ann Intern Med*. 2020;172:169-174. doi:10.7326/M18-3684
12. Holmgren AJ, Downing NL, Bates DW, et al. Assessment of electronic health record use between US and non-US health systems. *JAMA Intern Med*. 2020;14:251-259. doi:10.1001/jamainternmed.2020.7071
13. Downing NL, Bates DW, Longhurst CA. Physician burnout in the electronic health record era: are we ignoring the real cause? *Ann Intern Med*. 2018;8:50-51. doi:10.7326/M18-0139
14. O'Reilly KB. *E/M office-visit changes on track for 2021: what doctors must know*. American Medical Association. August 5, 2020. Accessed August 10, 2021. https://www.ama-assn.org/practice-management/cpt/em-office-visit-changes-track-2021-what-doctors-must-know.
15. American Medical Association. *Implementing CPT evaluation and management (E/M) revisions*. The American Medical Association. 2021. Accessed August 4, 2021. https://www.ama-assn.org/practice-management/cpt/implementing-cpt-evaluation-and-management-em-revisions.
16. Ong SY, Molly MJ, Brian W, O'Connell Ryan T, Richard G, Melnick Edward R. How a virtual scribe program improves physicians' EHR experience, documentation time, and note quality. *NEJM Catal*. 2021; 2(12). doi:10.1056/CAT.21.0294
17. Melnick ER, Dyrbye LN, Sinsky CA, et al. The association between perceived electronic health record usability and professional burnout among US physicians. *Mayo Clin Proc*. 2019;14:476-487. doi:10.1016/j.mayocp.2019.09.024
18. DiAngi YT, Stevens LA, Halpern-Felsher B, Pageler NM, Lee TC. Electronic health record (EHR) training program identifies a new tool to quantify the EHR time burden and improves providers' perceived control over their workload in the EHR. *JAMIA Open*. 2019;2(2): 222-230.
19. Hollister-Meadows L, Richesson RL, De Gagne J, Rawlins N. Association between evidence-based training and clinician proficiency in electronic health record use. *J Am Med Inform Assoc*. 2021;28(4):824-831.
20. Lourie EM, Utidjian LH, Ricci MF, Webster L, Young C, Grenfell SM. Reducing electronic health record-related burnout in providers through a personalized efficiency improvement program. *J Am Med Inform Assoc*. 2021;28(5):931-937.
21. Florig ST, Corby S, Devara T, Weiskopf NG, Mohan V, Gold JA. Medical record closure practices of physicians before and after the use of medical scribes. *JAMA*. 2022;1:1350-1352. doi:10.1001/jama.2022.13558
22. Apathy NC, Hare AJ, Fendrich S, Cross DA. Early changes in billing and notes after evaluation and management guideline change. *Ann Intern Med*. 2022;174(4):499-504. doi:10.7326/M21-4402
23. Nguyen OT, Turner K, Apathy NC, et al. Primary care physicians' electronic health record proficiency and efficiency behaviors and time interacting with electronic health records: a quantile regression analysis. *J Am Med Inform Assoc*. 2022;29:461-471. doi:10.1093/jamia/ocab272
24. Yan Q, Jiang Z, Harbin Z, Tolbert PH, Davies MG. Exploring the relationship between electronic health records and provider burnout: a systematic review. *J Am Med Inform Assoc*. 2021;28(5):1009-1021.
25. Tran B, Lenhart A, Ross R, Dorr DA. Burnout and EHR use among academic primary care physicians with varied clinical workloads. *AMIA Jt Summits Transl Sci Proc*. 2019;2019:136-144.
26. Micek MA, Arndt B, Tuan WJ, et al. Physician burnout and timing of electronic health record use. *ACI Open*. 2020;4(1):e1-e8.
27. Baxter SL, Apathy NC, Cross DA, Sinsky C, Hribar MR. Measures of electronic health record use in outpatient settings across vendors. *J Am Med Inform Assoc*. 2020;28:955-959. doi:10.1093/jamia/ocaa266
28. Melnick ER, Ong SY, Fong A, et al. Characterizing physician EHR use with vendor derived data: a feasibility study and cross-sectional analysis. *J Am Med Inform Assoc*. 2021;5:1383-1392. doi:10.1093/jamia/ocab011
29. Rotenstein LS, Jay Holmgren A, Lance Downing N, Bates DW. Differences in total and after-hours electronic health record time across ambulatory specialties. *JAMA Intern Med*. 2021;181:863-865. doi:10.1001/jamainternmed.2021.0256
30. Team RC. *R: A language and environment for statistical computing*. 2013. https://cran.microsoft.com/snapshot/2014-09-08/web/packages/dplR/vignettes/xdate-dplR.pdf.
31. RStudio Team. *R Studio*. 2015. http://www.rstudio.com/.
32. Bergé L. *Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm*. 2018. CREA Discussion Papers 13.

33. Grolemund G, Wickham H. Dates and times made easy with lubridate. *J Stat Softw*. 2011;40(3):1-25. https://www.jstatsoft.org/v40/i03/

34. Wickham H. The tidyverse. *R package ver* 1.1.1. 2017. https://slides.nyhackr.org/presentations/The-Tidyverse_Hadley-Wickham.pdf.

35. Holmgren AJ, Rotenstein L, Downing NL, Bates DW, Schulman K. Association between state-level malpractice environment and clinician electronic health record (EHR) time. *J Am Med Inform Assoc*. 2022;10:1069-1077. doi:10.1093/jamia/ocac034

36. Howe JL, Adams KT, Hettinger AZ, Ratwani RM. Electronic health record usability issues and potential contribution to patient harm. *JAMA*. 2018;319(12):1276-1278.

37. Melnick ER, Harry E, Sinsky CA, et al. Perceived electronic health record usability as a predictor of task load and burnout among US physicians: mediation analysis. *J Med Internet Res*. 2020;22(12):e23382.

38. 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.

39. Beasley JW, Wetterneck TB, Temte J, et al. Information chaos in primary care: implications for physician performance and patient safety. *J Am Board Fam Med*. 2011;24(6):745-751.

40. Sittig DF, Ash JS, Wright A, et al. How can we partner with electronic health record vendors on the complex journey to safer health care? *J Healthc Risk Manag*. 2020;40(2):34-43.

41. Colicchio TK, Cimino JJ. Clinicians' reasoning as reflected in electronic clinical note-entry and reading/retrieval: a systematic review and qualitative synthesis. *J Am Med Inform Assoc*. 2019;26(2):172-184.

42. Cimino JJ. Putting the "why" in "EHR": capturing and coding clinical cognition. *J Am Med Inform Assoc*. 2019;26(11):1379-1384.

43. Kahn D, Stewart E, Duncan M, et al. A prescription for note bloat: an effective Progress note template. *J Hosp Med*. 2018;13(6):378-382.

44. Ratwani RM, Reider J, Singh H. A decade of health information technology usability challenges and the path forward. *JAMA*. 2019;321(8):743-744.

45. Gomes KM, Ratwani RM. Evaluating improvements and shortcomings in clinician satisfaction with electronic health record usability. *JAMA Netw Open*. 2019;2(12):e1916651.

46. Adams KT, Pruitt Z, Kazi S, et al. Identifying health information technology usability issues contributing to medication errors across medication process stages. *J Patient Saf*. 2021;17(8):e988-e994.

47. McCoy AB, Russo EM, Johnson KB, et al. Clinician collaboration to improve clinical decision support: the Clickbusters initiative. *J Am Med Inform Assoc*. 2022;29(6):1050-1059.

48. Tajirian T, Jankowicz D, Lo B, et al. Tackling the burden of electronic health record use among physicians in a mental health setting: physician engagement strategy. *J Med Internet Res*. 2022;24(3):e32800.

49. Holmgren AJ, Apathy NC. Assessing the impact of patient access to clinical notes on clinician EHR documentation. *J Am Med Inform Assoc*. 2022;29:1733-1736. doi:10.1093/jamia/ocac120

50. Definitive Healthcare. *Top 10 ambulatory EHR vendors by 2019 market share*. Definitive Healthcare. 2019. Accessed August 4, 2021. https://www.definitivehc.com/blog/top-ambulatory-ehr-systems.

51. OCHIN. *Epic EHR—our members*. OCHIN. Accessed March 17, 2022. https://ochin.org/members.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.