AN EVALUATION OF AN ELECTRONIC MEDICAL RECORD (EMR) BASED SYSTEM TO CHARACTERIZE AND CORRELATE PHYSICIAN BURNOUT AND EMR USE

A Dissertation

Presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

KATIE WILKINSON

Dr. Lori Popejoy, Dissertation Supervisor

DECEMBER 2021

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

AN EVALUATION OF AN ELECTRONIC MEDICAL RECORD (EMR) BASED SYSTEM TO CHARACTERIZE AND CORRELATE PHYSICIAN BURNOUT AND EMR USE

presented by Katie Wilkinson

a candidate for the degree of doctor of philosophy

and hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Lori Popejoy

Dr. Robert Pierce, M.D.

Dr. Peter Tonellato

Dr. Dale Fitch

Dr. Lincoln Sheets

DEDICATION

To my husband Chad who has picked up my slack at home effortlessly and without a single word of complaint, you can now hunt and golf to your heart's content. And to my daughters, Eliana and Savannah who once asked me "Momma, why don't you just quit?" May you grow up understanding that perseverance pays off.

ACKNOWLEDGEMENTS

I want to acknowledge the Office of Clinician Well-Being at MU Health, Dr. Stephen Keithahn and Dr. Craig Rooney, who sponsored this work as one small part in their dedicated efforts to improve the clinician experience. I would like to acknowledge the engineers from the Tiger Institute Development at the Edge team who proposed and developed the SIBM in the EMR without whom this evaluation and future planned research would not be possible: Matt Tschannen, Alan Zentz, William Summers and Brad Rogers, and the engineering leader, Ben Naumann, and finally Dr. Cynthia Cipolla who collaborated throughout the project and did a masterful job of serving as my honest broker. I'd also like to acknowledge the research team that collaborated on the formulation and execution of this research: Dr. Peter Tonellato, Dr. Jung Kim, and Dr. Kangwon Seo.

I also want to thank my chair, Dr. Lori Popejoy who gave me the confidence and support to push this over the finish line. To my committee members, colleagues, and friends, Dr. Peter Tonellato, Dr. Dale Fitch, and Dr. Lincoln Sheets, I look forward to many research collaborations to come. Finally, I want to thank Dr. Robert Pierce who has collaborated on this work from the beginning: your clinical insights and informatics expertise are invaluable.

I'd also like to acknowledge clinical informatics leadership, especially Dr. Tom Selva, Dr. Bernard Eskridge, and Dr. Robert Pierce, as well as the small army of application, data and information technology folks in the Tiger Institute who work every day on the front line to improve healthcare and the provider experience. And finally, to my frontline clinical friends and colleagues for all your support and feedback.

Table of Contents

ACKNOWLEDGEMENTS ii

List of Tables vi

List of Illustrations viii

List of Abbreviations ix

Abstract xi

Chapter 1: General Introduction 1

Chapter 2: Current State of the Literature 7

2.1 Background 8

2.1.1 Measuring Burnout 10

2.2 Materials and Methods 12

2.3 Results 13

2.3.1. Physician Characteristics and Burnout 14

2.3.2 EMR Use and Burnout 15

2.3.3 Measurement of EMR Use and Burnout 17

2.4 Discussion 21

2.5 Conclusion 23

Chapter 3: Research Proposal 24

3.1 Specific AIMs 24

3.2.1 Significance 26

3.3 Innovation 30

3.4 Approach: 31

3.4.1 Background 31

3.4.2 Conceptual Framework 32

3.4.3 Scientifically Relevant Preliminary Studies: 32

3.4.4: Specific Aims 37

Chapter 4: Single-item burnout measure in the clinical workflow 46

4.1 Abstract 46

4.1 Introduction 47

4.1.1 Ethical considerations 48

4.2 Materials and Methods 48

4.2.1 SIBM Instrument and pilot 48

4.2.2 Survey Administration 49

4.3 Results ... 51

4.3.1 Impact of Population Characteristics and Survey Administration on Response rates ... 51

4.3.3 Impact of individual characteristics on burnout response scores ... 53

4.3.4 Impact of survey administration on burnout scores ... 54

4.3.4 Test-retest reliability ... 55

4.3.5: Thematic Analysis of text responses ... 56

4.4 Discussion ... 57

4.5 Limitations ... 59

4.6 Conclusions ... 59

Chapter 5: Characterization of EMR usage patterns by specialty and venue ... 60

5.1 Introduction ... 61

5.2 Materials and Methods ... 62

5.2.1 A Note on Outliers ... 65

5.3 Results ... 66

5.3.1 Physician EMR usage month over month and between venues ... 66

5.3.2 Factor analysis of EMR use metrics ... 68

5.4 Discussion: ... 74

5.5 Limitations ... 78

5.6 Conclusion ... 79

Chapter 6: Single population correlation of EMR metrics to SIBM ... 80

6.1 Introduction and Background: ... 82

6.2 Materials and Methods: ... 83

6.2.1 Data Preparation ... 85

6.2.2 Regression analysis ... 86

6.3 Results ... 87

6.4 Discussion ... 91

6.5 Limitations ... 94

6.6 Conclusion: ... 95

Chapter 7: Conclusion ... 97

7.1 Dissertation Overview and Findings ... 97

7.1.1 Approach ... 98

7.1.2 Major Findings ... 99

7.2 Discussion and Future Directions 101
APPENDICES 103
Appendix A: Single Item Burnout Measures Cited 103
Appendix B: Audit log variables 104
Appendix C: Full Factor Analysis Results 105
Appendix D: Specialty Groupings 108
Appendix E: Detailed Intra-Class Correlation Results for all models 109
BIBLIOGRAPHY 110
VITA 128

List of Tables

Table 2. 1: Summary of study characteristics and measurement approaches[10] 19

Table 3. 1: Helpful acronyms and their corresponding terms 31

Table 3. 2: Prescribing providers captured in Cerner Advance event logging data....35

Table 3. 3: Proposed high-level EMR metrics (See Appendix A for data dictionary)....36

Table 3. 4: Individual physician and survey metadata 38

Table 4. 1: Survey and provider characteristics 51

Table 4. 2: Total physicians surveyed by specialty and demographics 51

Table 4. 3: Response rates by specialty and demographics 52

Table 4. 4: Response Rates relative to time, day of week and month of survey....53

Table 4. 5: Burnout scores relative to individual provider specialty and demographics....54

Table 4. 6: Relationship of survey characteristics and burnout responses....55

Table 4. 7: Month to month change in provider burnout responses....56

Table 5. 1: Specialty Groupings 63

Table 5. 2: Counts of physician-months in each specialty by venue 66

Table 5. 3: Factor loadings exceeding a .3 cutoff for all venues...................69

Table 5. 4: Regression of all venue factors on flag for ambulatory and inpatient physicians 70

Table 5. 5: Regression of overall EMR use factors on specialty 71

Table 5. 6: Significant relationships of all venue factors with specialty groups...71

Table 5. 7: Factor loadings exceeding a .3 cutoff for ambulatory venue......72

Table 5. 8: Regression of ambulatory EMR use factors on specialty 72

Table 5. 9: Significant relationships of ambulatory venue factors with specialty groups 72

Table 5. 10: Factor loadings exceeding a .3 cutoff for inpatient venue 73

Table 5. 11: Regression inpatient EMR use factors on specialty..................74

Table 5. 12: Significant relationships of inpatient venue factors with specialty groups 74

Table 6. 1: Univariate regression results overall and by Venue....................89

Table 6. 2: Multivariate regressions of individual EMR usage variables controlling for gender, age and specialty 90

List of Illustrations

Figure 2. 1: Summary of articles included in review... 14

Figure 3. 1: Conceptual model for research design 33

Figure 3. 2: SIBM as distributed through the EMR Message Center. 34

Figure 3. 3: Pilot survey responses 34

Figure 3. 4: SIBM response rates and average scores collected in repeated survey of all attending physicians at an AMC over a period of three months. 35

Figure 3. 5: Preliminary data illustrating percent of time spent in ambulatory vs inpatient based on time spent documenting on ambulatory or inpatient encounters respectively... 41

Figure 4. 1: The SIBM delivered through the EMR Message Center 49

Figure 5. 1: Month to month ICC for all venues... 67

Figure 5. 2: Month over month ICC: Single venue vs inpatient to ambulatory................ 68

List of Abbreviations

AMC	Academic Medical Center
CEHRT	Certified Electronic Health Record Technology
CDS	Clinical Decision Support
CHIME	College of Healthcare Information Management Executives
CMIO	Chief Medical Information Officer
CPOE	Computerized Physician Order Entry
EFA	Exploratory Factor Analysis
EHR	Electronic Health Record: Equivalent to EMR for purposes of this paper.
EMR	Electronic Medical Record
HIMSS	Health Information and Management Systems Society
HIT	Healthcare Information Technology
HITECH	Health Information Technology for Economic and Clinical Health
ICD-11	11th Revision of the International Classification of Diseases
JAMIA	Journal of the American Medical Informatics Association
KMI	Kaiser-Meyer Olkin
LHS	Learning Health System
MBI	Maslach Burnout Inventory
MBI-HSS	22-item MBI-Human Services Survey
MBI-GS	16-item MBI-General Survey
ML	Machine Learning
OCW	Office of Clinical Well-Being
PFI	Professional Fulfillment Index

SIBM	Single-Item Burnout Measure
SQBS	Single-Question Burnout Score (Mini-Z)
SUS	System Usability Scale
TLI	Tucker Lewis Index

AN EVALUATION OF AN ELECTRONIC MEDICAL
RECORD (EMR) BASED SYSTEM TO CHARACTERIZE
AND CORRELATE PHYSICIAN BURNOUT AND EMR USE

Katie Wilkinson

Dr. Lori Popejoy, Dissertation Supervisor

ABSTRACT

Burnout disproportionately affects healthcare workers and continues to rise, contributing to cost, quality, and patient safety risk in an already overburdened United States healthcare system. While the causes of burnout are complex, evidence suggests that Electronic Medical Record use (EMR) is one major contributor due to the increased clerical burden that decreases patient contact time and disrupts the provider clinical workflow. The challenge of improving the physician EMR experience is exacerbated both by variability across venues and specialty. Targeted training and optimization efforts are generally deployed one-time at a clinic or specialty level but are challenging to deploy longitudinally and in surveillance mode due to the cost and effort of administering traditional survey instruments. To address this challenge, we deployed a single-item burnout measure (SIBM) at the University of Missouri Healthcare, an academic medical center (AMC), to test the feasibility and reliability of capturing longitudinal physician self-reported burnout through the EMR. We further evaluated the utility of the proposed EMR event logging data to discriminate presumed differences in workflow between venues (inpatient, outpatient, and emergency department) and specialty groups (primary care, surgical, non-surgical medical, and emergency) and then correlated the EMR data with the burnout data to demonstrate how this EMR-based longitudinal platform can be used to understand how varying EMR use correlated to burnout.

Chapter 1: General Introduction

Unlike other mental health disorders, the general definition of burnout as a workplace-related phenomenon resulting from poorly managed chronic workplace stress was codified in the 11^{th} Revision of the International Classification of Diseases in 2019 (ICD-11) (World Health Organization, 2021). The prevalence of burnout in healthcare providers, as well as the impact of burnout on individuals and the healthcare system, already well documented, was only exacerbated by the worldwide COVID pandemic and the accompanying isolation, high-acuity patient load, and staffing shortages. The causes of burnout are multi-faceted including individual, work, and organization-related factors. With the exponential rise of EMR use, characterized by an increase of documentation burden associated with regulatory, quality, and patient safety requirements levied on healthcare providers, the EMR, and Health Information Technology (HIT) more broadly, is often identified as an area for improvement.

MU Health has a long history of focus on improving the provider EMR experience beginning in 2006 with precocious HIT-backed population management efforts in Family & Community Medicine (FCM). These efforts led to a major framework improvement in the EMR and culminated in an HIT-centric strategic public-private partnership with Cerner Corporation, an HIT company out of Kansas City, MO. The success of these improvement efforts resulted in changes such as on-site support from application and training teams who conducted regular rounding in clinical areas, specialty-specific optimization, adoption of the essential clinical dataset (Karp, Freeman, Simpson, & Simpson, 2019), engaged clinical leadership and vetted governance. These improvements resulted in Health Information and Management Systems Society

(HIMSS) level 7 recognition in inpatient and ambulatory by 2013. Moreover, MU Health has had 11 years of recognition as a "Most Wired" healthcare organization by the College of Healthcare Information Management Executives (CHIME), including level nine attainment in 2021. In 2021 MU Health was awarded the distinction of being one of only nine healthcare organizations worldwide to win the HIMSS Davies award for a second time for use of HIT to substantially improve clinical care delivery, patient outcomes, and population health around the world.

In parallel to this activity, MU Health established the Office of Clinician Well-Being (OCW) to lead efforts and provide resources to combat clinician burnout. Clinician feedback on the EMR was first captured by two cumbersome annual EMR satisfaction survey targeted at nurses and physicians and later through a 3rd-party survey which brought the added benefit of benchmarks to other organizations. Similarly, the OCW collected feedback on physician burnout through the Mini-Z administered by a third-party. While this feedback was valuable, there were two major gaps. First, because of the survey administration, anonymization, and aggregation of the results, it was not possible to correlate the individual-level feedback back to any specific cause. Second, even if it were possible to do so, the surveys were too cumbersome to support the longitudinal measurement of clinician burnout over time needed to support measuring the impact of improvement interventions. This phenomenon is not unique to MU Health. In their 2019 systematic review of Interventions to improve the use of EMRs in primary healthcare, Hamade et al. found only 12 measured interventions, and 11 of those measured success through a process measure such as use of a particular tool, or completeness of documentation. Only one looked at improvement of data quality markers over time and none measured any impact on the clinician well-being (Hamade, Terry, & Malvankar-Mehta, 2019). Similarly, in another 2021 systematic review specifically focused on EMR use and burnout, Yan et al. found that just 3 out of 26 articles incorporated any pre-post intervention measurement of the impact that improvements had on physician well-being.

Nature and purpose of the study

In 2019 a cross-functional group consisting of leaders from the OCW, an onsite engineering innovation team, physician informatics leadership, and researchers formed to discuss ways to identify and improve HIT contributions to burnout. The immediate goal for the OCW was to see if it is possible to capture longitudinal burnout measures within the EMR to measure the impact of improvement efforts. For the research team, the potential availability of a de-identified, individual provider-level measure of burnout raised the possibility not just of measuring the impact of interventions, but also of potentially predicting those behaviors that are likely to result in burnout to target just-in-time interventions to prevent burnout. With this future vision in mind, the team embarked on this work to demonstrate the ability to measure physician burnout and correlate it to EMR usage patterns through the development of an entirely EMR-based measurement system.

Scientific approach

Demonstration of the feasibility and reliability of an EMR-based measurement system to understand how the EMR contributes to burnout depends on two things. First, a reliable measure of burnout. Second, a reliable way to measure EMR usage and clinical workflow. Since we knew of no existing EMR-based measure of burnout, the OCW team first piloted and later deployed an externally validated Single-Item Burnout Measure in the EMR. The surveys were distributed, and responses collected through the EMR's e-mail-like inbox functionality, "message center". To understand the feasibility and reliability of this repeated measure, the survey was deployed to randomized day of week and time of day cohorts of all attending physicians at MU Health and evaluated for the impact of individual level factors and survey administration factors on both response rates and burnout rates. Results were evaluated for any apparent bias that would invalidate this approach for capturing repeated measures of burnout.

Characterization of EMR usage is somewhat more problematic. Due in a large part to regulatory requirements to capture specific information about any access to protected health information, detailed event logs (also referred to as audit logs or usage logs in the literature) are provided through every commercially available EMR in the United States. Major vendors such as Epic and Cerner provide these data back to clients for purposes of quality improvement. These data capture information on time spent on defined activities, use of decision support and advanced features, and application responsiveness and stability, all of which characterize variations in the overall physician EMR experience. At the source, the data constitute very big data with high fidelity and present many possibilities for future longitudinal understanding of clinical workflow. However, due to the size and complexity of those data, the current study was restricted only to pre-processed data provided by the vendor by month, physician and, for time in EMR measures, venue of practice (ambulatory, inpatient or emergency department). Because of the complexity and high-dimensionality of the EMR data provided by the vendor, we used exploratory data mining techniques to first identify underlying patterns of EMR usage, both overall and by sub-groups, and looked at stability of EMR usage over time and between venues. Guided by this exploratory analysis, we then used the event log data and burnout scores to identify relationships between EMR usage and burnout both with and without controls for individual-level characteristics.

Significance of the Study

The primary significance of the study is the demonstration of the feasibility and reliability of collecting repeated measures of burnout within the EMR. This is expected to aid in the understanding of burnout as a longitudinal phenomenon, but also to begin to directly measure within a prospective study design how we can combat burnout. This study further takes the step of demonstrating the utility of coupling this data with existing EMR event log data to deploy an ecosystem for understanding and improving EMR-related causes of burnout.

Conclusion

The successful execution of this study design represents a first step. It was expected that the aggregated nature of the event log data would limit our ability to glean meaningful insights into the clinical workflow. The exploratory analysis produced interesting insights into EMR usage between venues, but much work is needed and already underway in the research community to continue to understand meaningful patterns of EMR usage and clinical workflow. Despite the limitation, we were able to demonstrate the successful deployment of the survey instrument in the EMR to collect repeated measures and this, coupled with event log data, brought to light evidence of how

EMR usage relates to burnout and provides insight into the possibilities for future research.

Chapter 2: Current State of the Literature

Background: The rapid adoption of Electronic Medical Records (EMRs) has been frequently cited as a contributor to burnout. This rapid adoption also presents an opportunity to use data to in a Learning Health System (LHS) to target improvement opportunities and measure impact.

Objective: To synthesize the current evidence on EMR contributions to physician well-being and approaches that have been used to measure EMR activity and burnout.

Methods: Source manuscripts and associated evidence from two applicable systematic reviews published in May 2021 were reviewed. Evidence from the 52 articles was synthesized to evaluate the impact of individual-level characteristics and EMR use on Physician well-being and the relationship between the two. Articles were further evaluated for measurable EMR use and measurable burnout and any articles that were not peer-reviewed or focused exclusively on residents and/or fellows were excluded. The remaining 27 articles were reviewed to define their approach to measuring EMR association with burnout. Articles were coded for sources of data defining EMR Use, burnout, study type and methods.

Results: Evidence synthesis shows after-hours time, perceptions of EMR, and usability are predictive of physician well-being while total time in EMR and on activities such as chart review are inconclusive. Evidence was found of the association between burnout and specific functionality such as Computerized Physician Order Entry and Message Center volume and alerts. 26% of the articles (7/27) used EMR-derived data to measure EMR use, 70% (19/27) used survey instruments, 4% (1/27) used a focus group supplemented with administrative data and 4% (1/27) measured EMR use at the

organizational level.   A Single Item Burnout Measure (SIBM) as the primary measure of burnout was used in 48% (n =13) of studies, but all 27 involved additional survey measures or focus group information. Only 11% (3/27) measured any pre-post intervention while the rest were cross-sectional and observational only.

Conclusion: While there is a great deal of evidence on the link between EMR use and burnout, this evidence relies on varied and often burdensome survey instruments which do not lend themselves continuous measurement and improvement in a Learning Health System. An increase in the use of EMR-derived data in the last 2 years and validated use of an SIBM may provide opportunity for development of a data-driven LHS model of EMR improvement.

2.1 Background

Burnout as a psychological syndrome was first defined in the 1970s by Herbert J. Freudenberg after he personally experienced a feeling of physical and emotional exhaustion associated with his work with the Free Clinic Movement. (Freudenberger, 1974) Maslach and colleagues later defined burnout as a syndrome characterized by emotional exhaustion or loss of enthusiasm for work, depersonalization or cynicism, and a lack of personal accomplishment (Maslach, Jackson, Leiter, Schaufeli, & Schwab, 1986) (Maslach, Schaufeli, & Leiter, 2001). Since that time, it is estimated that as many as 1,000 articles are published each year on Burnout syndrome (Maslach & Leiter, 2014) (Schaufeli, Leiter, & Maslach, 2010).

A large body of literature has established the prevalence of burnout among physicians. In one systematic review, Rotenstein and colleagues identified 182 articles quantifying the prevalence of physician burnout as anywhere between 0 and 85%

(Rotenstein et al., 2018), while another meta-analysis calculated the mean burnout rate to be 44% (West, Dyrbye, Erwin, & Shanafelt, 2016). Shanafelt et al. found that burnout is significantly more prevalent among US nurses and physicians than other professions (Tait D. Shanafelt et al., 2012). Notably, they found that while an MD or DO is more likely to experience burnout compared to a high school graduate, individuals with other college degrees were less likely to experience burnout compared with high school graduates with odds ratios decreasing with higher degree attainment . The underlying causes of burnout identified include: a) individual factors such as age, gender, marital status, specialty, and job position; b) work factors such as workload, shift work, administrative duties, team cohesion, and c) organizational factors such as negative leadership, adequate rewards, and workload expectations (Amoafo, Hanbali, Patel, & Singh, 2014; Azam, Khan, & Alam, 2017; Patel, Bachu, Adikey, Malik, & Shah, 2018). Studies have long suggested that the growth in EMRs contributed to physician burnout by adding to clerical burden and depersonalizing medicine by reducing patient interaction. In a highly cited time-motion study of EMR activity across four specialties, Sinsky et al. (2016) found physicians spend nearly two additional hours on clerical work (EMR and desk time) for every hour of patient contact (C. Sinsky et al., 2016). This finding was reinforced by Arndt et al. in Primary Care using event log data validated to time-motion observations (Arndt et al., 2017). While the initial shift to EMRs may have represented more of a like for like move from paper to electronic, subsequent clerical burden has been driven factors such as increased regulatory expectations for documentation, patient safety and quality measure tracking, and new expectations of patient-provider interaction (e.g., through message center) which drives high-level metrics like patient satisfaction (Patel et al., 2018; Rossetti et al., 2021).

The impacts of physician burnout include job dissatisfaction, increased sick leave, turnover, lower productivity (Halbesleben & Rathert, 2008; Tait D. Shanafelt et al., 2012; Soler et al., 2008; Zhang, Gunter, Liebovitz, Tian, & Malin, 2011) (C. S. Dewa, Loong, Bonato, Thanh, & Jacobs, 2014) (Hoff, Whitcomb, & Nelson, 2002)} and lower perceived ability to do their job (Ruitenburg, Frings-Dresen, & Sluiter, 2012). Halbesleben & Rathert (2008) identified a link between burnout and lower patient satisfaction and adverse patient outcomes in terms of length of stay (Halbesleben & Rathert, 2008). Panagioti et al. conducted a meta-analysis identifying that burnout was associated with two-fold increase odds for unsafe care, unprofessional behaviors, and low patient satisfaction (Panagioti et al., 2018). Dewa et al. estimated the financial impact of increased sick leave and early retirement to be $213.1 million (CAN) (Carolyn S Dewa, Jacobs, Thanh, & Loong, 2014) while Hans et al. estimate the overall cost to the healthcare system to be $4.6 billion (US) annually due to lost clinical revenue and the cost of turnover (Han et al., 2019).

2.1.1 Measuring Burnout

The gold standard for measuring burnout has long been the Maslach Burnout Inventory (MBI), a 22-item instrument that measures burnout along the three identified dimensions: emotional exhaustion, depersonalization, and personal accomplishment (Maslach et al., 1986; Maslach, Jackson, Leiter, Schaufeli, & Schwab, 2016). One systematic review (Rotenstein et al., 2018) of studies of burnout in attending physicians found that 85.7% of studies (156 of 182) used some version or sub-scale of the MBI, including the full-length 22-item MBI-Human Services Survey (MBI-HSS), the 16-item MBI-General Survey (MBI-GS), or a sub-section of the MBI-HSS including single-item burnout measures demonstrated to validate against emotional exhaustion and depersonalization (West et al., 2016). The remaining studies mostly leveraged public domain methods to measure burnout including the 16-item Astudillo and Mendinueta Burnout Questionnaire (Astudillo & Mendinueta, 1996), the 54-item Modified Compassion Satisfaction and Fatigue Test (Weintraub, Geithner, Stroustrup, & Waldman, 2016), the 19-item Copenhagen Burnout Inventory (Kristensen, Borritz, Villadsen, & Christensen, 2005), the 40-item Hamburg Burnout Inventory (Burisch, 1984), the Pines and Aronson Burnout Measure (Malakh-Pines, Aronson, & Kafry, 1981), the 20-item Spanish-language Questionnaire for the Evaluation of Work-Related Burnout Syndrome (CESQT) (Gil-Monte, Rojas, & Ocaña, 2009), the 10-item Zero Burnout Program Survey (Shimotsu, Poplau, & Linzer, 2015), and various single-item measures of self-perceived burnout, including the measure of Rohland et al. (Rohland, Kruse, & Rohrer, 2004). The authors had planned a meta-analytic pooling of study results but were unable to do so primarily because of the heterogeneity in study design, burnout instruments, and cutoffs.

More recently the 16-Item Professional Fulfillment Index (PFI) was introduced as a free alternative to the MBI (Trockel et al., 2018). Using principal component analysis, Trockel et al. demonstrated that the PFI correlated well to corresponding MBI questions, but also captured professional fulfilment, work exhaustion and interpersonal disengagement as captured by various validated instruments. In 2019 Olson et al. validated the Mini-Z as a brief alternative to identify stressors associated with burnout and guide interventions (Olson, Marchalik, et al., 2019; Olson, Sinsky, et al., 2019). As an alternative to the lengthier full MBI, West et al. first demonstrated that the single “I feel burned out from my work” question rated on a seven-point scale from “Never” to “Every day” best validated to the emotional exhaustion domain of the MBI (West, Dyrbye, Sloan, & Shanafelt, 2009). Since that time, multiple versions of single-item burnout measures have been offered as externally validated simple alternatives to costly or lengthy alternatives such as the MBI (Dolan et al., 2015; Olson, Sinsky, et al., 2019; Rohland et al., 2004)

2.2 Materials and Methods

There are many rigorous and recent systematic reviews providing evidence on adoption of EMRs; prevalence of EMR use; causes and consequences of burnout and the effect on productivity; characterizing EMR use; defining documentation burden; and laying out interventions to improve EMR use (C. S. Dewa et al., 2014; Dutta & Hwang, 2020; Hamade et al., 2019; Patel et al., 2018; Pinevich, Clark, Harrison, Pickering, & Herasevich, 2021; Rotenstein et al., 2018; West et al., 2016; Moy et al., 2021). Two systematic reviews published in the May 2021 special burnout edition of the Journal of the American Medical Informatics Association (JAMIA) focused specifically on the link between Healthcare IT and physician well-being. Both systematic reviews included the concepts of Health IT and physician wellness and articles focused on providers, but because of variations in databases, search terms and approach, it is useful to synthesize findings from both. In the first, Yan et al. looked explicitly at research between burnout and EMR use (Yan, Jiang, Harbin, Tolbert, & Davies, 2021). There is significant overlap between the Yan and Nguyen articles: 9 of the 35 articles in Nguyen et al. (26%) overlap with Yan et al. While 13 measured burnout directly, the Nguyen et al. review also assessed organizational, physician and IT-related factors associated with EMR-related impacts on well-being. This broader approach in their search strategy adds to the evidence provided by Yan et al. by allowing indirect EMR-related measures of burnout including: efficiency and resources; workload and job demands; work-life integration; and organizational culture and value. (Nguyen et al., 2021).

Because of the recency and relevance of the available evidence, this review will not attempt to reproduce these systematic reviews. Instead, the existing search strategies, codified data, and source manuscripts will be leveraged in two ways. First, a summary of the evidence of the link between EMRs and burnout, including more general physician well-being, using all source manuscripts will be provided. Second, a subset of articles that specifically measure EMR use and the concept of burnout were selected. Those manuscripts that were not peer-reviewed or only focused on residents and physicians were excluded from the analysis. These articles were coded according to how they measured EMR use and burnout, study design, number of respondents and response rate. Where available, coding was validated against the source systematic review for concordance. If information in the source systematic review couldn't be verified, it was excluded. I subsequently analyzed and summarized the approaches to measuring the link between EMR use and burnout.

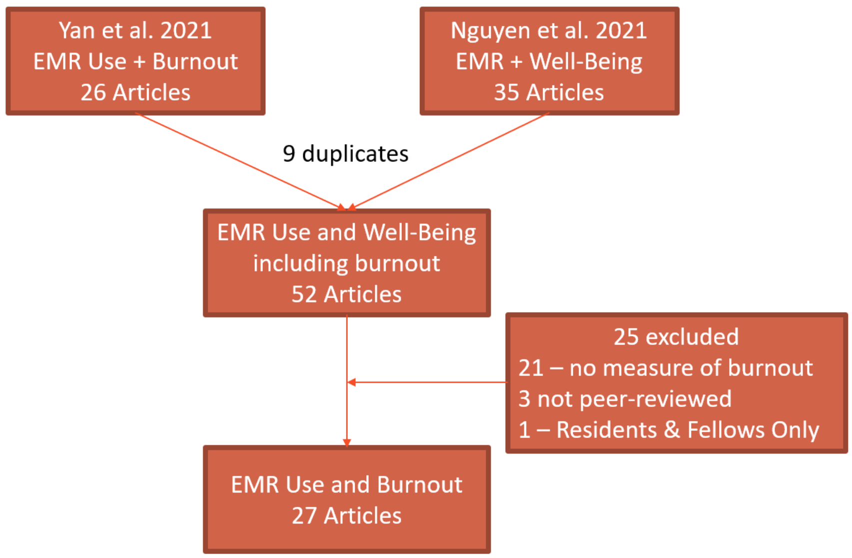
2.3 Results

Figure 2.1 shows the articles included. In total between the two reviews there were 52 unique manuscripts identified that evaluate the link between EMR use and physician well-being (61 total with 9 duplicates). An additional 25 were subsequently excluded because they either had no objective measure of burnout, were not peer-reviewed, or only included residents and fellows.

2.3.1. Physician Characteristics and Burnout

Many individual physician characteristics have been shown to correlate to burnout. Three confounding variables are controlled for in this study and are therefore

Figure 2. 1: Summary of articles included in review



the focus here: gender, age, and specialty. Studies about the effect of female gender on burnout were mixed. Five studies linked female gender with more time in the EMR and higher burnout (Gardner et al., 2019; Gupta, Murray, Sarkar, Mourad, & Adler-Milstein, 2019; Kroth et al., 2019; E. R. Melnick et al., 2020; T. D. Shanafelt et al., 2016) yet eight others found no association between burnout related factors and gender (Butler & Johnson, 2016; Gardner et al., 2019; E. R. Melnick et al., 2020; Robertson, Robinson, & Reid, 2017; Ward, 2019; Kroth et al., 2019; Mehta et al., 2019). Most studies did not find a direct association between age and burnout for attending providers (Gardner et al., 2019; Kroth et al., 2019; Mehta et al., 2019; E. R. Melnick et al., 2020), but older age was found to be associated with lower usability perception of EMR ease of use, perceived physician productivity, and overall EMR satisfaction (Butler & Johnson, 2016; E. R. Melnick et al., 2020; Wylie, Baier, & Gardner, 2014). Many of the studies either focused on a single specialty or lacked adequate numbers of participants to assess the impact of specialty on EMR causes of burnout. Where specialty was considered relative to clerical burden, variations were noted between specialties. For example, Shanafelt et al. (2016) found that family medicine and emergency medicine experienced high burnout and were less likely to think clerical burden related to patient care was reasonable as compared with Internal Medicine. Pediatrics, Surgical and OB/Gyn specialties experienced less burnout than Internal Medicine. Similarly, Hilliard et al. (2020) found that Primary Care Physicians (PCPs) were more likely both to have higher clerical burden and burnout levels. Two studies reported mixed results on EMR usability with Melnick et al. (2020) finding that General Internal Medicine rated EMR usability higher than other specialties (Butler & Johnson, 2016). Gardner et al. (2019) reported the highest HIT-related stress in primary care-oriented specialties. By contrast, Robertson, Robinson, & Reid (2017) noted no relationship of the specific primary care specialty to HIT-related burnout and noted no difference in EMR-related burnout between Primary Care, Procedural and Non-Procedural specialties (Kroth et al., 2019; Saag, Shah, Jones, Testa, & Horwitz, 2019). Two studies found that generalists have more inbox volume than specialists (Tai-Seale et al., 2019) (Murphy et al., 2016)

2.3.2 EMR Use and Burnout

The reviews uncovered mixed results in the studies looking at Objective EMR use. For example, total time in the EMR was found to be associated with emotional exhaustion and depersonalization as measured by the MBI-HSS (Domaney, Torous, &

Greenberg, 2018), while another study found that greater time in the EMR was associated with higher satisfaction with the interface although this effect was mitigated by the physician's age (Khairat et al., 2018) . While six studies found that time on EMR at home or after work was associated with increased burnout, EMR frustration levels, and decreased work-life balance (Gardner et al., 2019; Robertson, Robinson, & Reid, 2017; Privitera et al., 2018; Hauer, Waukau, & Welch, 2018; Olson, Sinsky, et al., 2019; Tai-Seale et al., 2019), another four found no association (Gilleland et al., 2014; Harris, Haskell, Cooper, Crouse, & Gardner, 2018; Mehta et al., 2019; Rassolian et al., 2017). Micek et al. found that EMR time on weekends and holidays increased burnout while time spent after-hours on weeknights or administrative days did not increase burnout (Micek et al., 2020). Three other studies found that general user proficiency and efficiency measures were not associated with burnout (Adler-Milstein, Zhao, Willard-Grace, Knox, & Grumbach, 2020; Olson, Sinsky, et al., 2019; Rassolian et al., 2017). Perception of usability, as measured by the System Usability Scale (SUS); and intuitive interfaces and provider belief that the EMR provided clear prompts; were associated with lower levels of perceived effort and lower burnout (Khairat et al., 2018) (Burke et al., 2017; E. R. Melnick et al., 2020), while Kroth found no association between user difficulty and burnout (Kroth et al., 2019). A high number of message alerts and inadequate time to deal with them was demonstrated to be associated with wellness in five studies (Adler-Milstein, Zhao, Willard-Grace, Knox, & Grumbach, 2020; Gregory, Russo, & Singh, 2017; Hilliard, Haskell, & Gardner, 2020; Tai-Seale et al., 2019; Friedberg et al., 2019) while the time on inbox alert messages, number of result messages and patient portal usage were not associated with burnout (Gregory et al., 2017; Hilliard et al., 2020; T. D. Shanafelt et al., 2016). One study found CPOE use was associated with burnout (T. D. Shanafelt et al., 2016) while another found the number of orders and use of preferences lists or "Smart Sets" was not associated with burnout (Hilliard et al., 2020). Time spent on EMR documentation and the lack of sufficient time to complete documentation were consistently associated with burnout in the seven studies that examined them (Domaney et al., 2018; Gardner et al., 2019; Olson, Sinsky, et al., 2019; Harris et al., 2018; Mehta et al., 2019; Privitera et al., 2018; Rassolian et al., 2017), but two studies found that chart review specifically was inconclusive (Domaney et al., 2018; Hilliard et al., 2020). Basic EMR functions, including tools to gain efficiency in EMRs such as Order Sets, pre-charting, Chart Search Function, and the ability to write EMR notes while with patients were not found to be associated with burnout. The only EMR efficiency function found to be associated was the use of "copy and paste" which was found to be negatively associated with burnout (Domaney et al., 2018; Adler-Milstein, Zhao, Willard-Grace, Knox, & Grumbach, 2020; Gardner et al., 2019; Harris et al., 2018; Menachemi, Powers, Au, & Brooks, 2010). There is strong evidence that negative perceptions of EMRs such as disagreement that EMRs keep people safe or improve efficiency or patient care, and dissatisfaction with clerical tasks, are negatively associated with burnout across all studies. (Gardner et al., 2019; Harris et al., 2018; Khairat et al., 2018; Marckini, Samuel, Parker, & Cook, 2019; Somerson, Patton, Ahmed, Ramey, & Holliday, 2020).

2.3.3 Measurement of EMR Use and Burnout

Table 2.1 provides a summary of study characteristics and measurement approaches for the literature on EMR use and burnout. The earliest article found to fit the inclusion criteria was 2014 with 56% (n=15) articles published in the last 2 years. There were a variety of burnout scales used, 33% (n=9) use some form or sub-scale of the MBI while 15% (n=4) leveraged the Mini-Z, 4% (n=1) used the Shirom-Melamed Burnout Questionnaire, and 48% (n=13) used a version of an SIBM. The two most common SIBMs are the measure cited by Rohland et al. (26% n=7) and the single question from the Mini-Z (15% n=4) with one additional proprietary SIBM and one adapted from the "I feel burned out from my work" question in the emotional exhaustion sub-scale of the MBI. All but one survey included at least one single-item direct measure of burnout (See Appendix A in supplementary material for all single-item burnout measures). Except for Gilleland et al. (2014), 100% of the SIBM studies included additional survey items to supplement their measurement of burnout with wellness concepts such as work-life balance, professional fulfillment, emotional exhaustion, and stress.

Of the 27 articles surveyed, 26% (n=7) used EMR-Derived data to measure EMR use. Apart from one 2014 study, all the articles using EMR-derived data were published between 2019 and 2020 and the EMR data used by the 2014 study was limited to login/logoff times (Gilleland et al., 2014), unlike later studies that leveraged broad EMR data. Surveys to capture information about EMR use or perceptions of EMRs were used in 63% (n=17) of studies. It is noteworthy that four of those leveraged information from the Mini-Z, a survey instrument that captures burnout in addition to stress and satisfaction as outcomes and documentation time pressure, EMR use at home and EMR proficiency as three of the seven drivers of those outcomes, along with work control, work chaos, teamwork, and leadership value alignment. Three of the four studies used additional survey measures of EMR use while one (Rassolian et al. 2017) relied solely on the

Table 2. 1: Summary of study characteristics and measurement approaches[10]

Article	Year	Measure of EMR Use	EMR	Study Design	Burnout Measure	Respondents	Response Rate
Babbott et al.	2014	Survey	NA	Cross-Sectional	SIBM - Rohland	379	50%
Contratto et al.	2017	Focus Group[1]	NA	Pre-Post Intervention[4]	SIBM - MBI[7]	7	100%
Domaney et al.	2018	Survey	NA	Cross-Sectional	MBI-HSS	52	1%
Gardner et al.	2019	Survey	NA	Cross-Sectional	SIBM - Mini-Z	1792	43%
Melnick et al.	2020	Survey	NA	Cross-Sectional	MBI-HSS	870	70%
Olson et al.	2019	Survey	NA	Cross-Sectional	MBI-HSS; Mini-Z	475	38%
Robertson et al.	2017	Survey	NA	Cross-Sectional	SIBM - Rohland	585	68%
Shanafelt et al.	2016	Survey	NA	Cross-Sectional	MBI-HSS	6560	18%
Tai-Seale et al.	2019	EMR-Derived	Epic	Cross-Sectional	SIBM - Rohland	934	72%
Adler-Milstein et al.	2020	EMR-Derived	Epic	Cross-Sectional	MBI-GS[8]	87	67%
Giess et al.	2020	Survey	NA	Cross-Sectional	SIBM - Rohland	159	78%
Gilleland et al.	2014	EMR-Derived[2]	Centricity	Cross-Sectional	SIBM - Proprietary	139	NR
Gregory et al.	2017	Survey	NA	Cross-Sectional	Shirom-Melamed	16	NR
Harris et al.	2018	Survey	NA	Cross-Sectional	SIBM - Mini-Z	371	31%
Hilliard et al.	2020	EMR-Derived	Epic	Cross-Sectional	SIBM - Mini-Z	422	7%
Kroth et al.	2019	Survey	NA	Cross-Sectional	SIBM - Rohland	282	44%
Marckini et al.	2019	Survey	NA	Cross-Sectional	MBI-HSS	110	29%
Mehta et al.	2019	Survey	NA	Cross-Sectional	Mini-Z	2274	21%
Pozdnyakova et al.	2018	Manual Logs	Epic	Pre-Post Intervention[5]	SIBM - Rohland	6	100%
Privitera et al.	2018	Survey	NA	Cross-Sectional	Mini-Z	1048	4%
Sieja et al.	2019	Survey	NA	Pre-Post Intervention[6]	MBI-HSS[9]	113	55%
Somerson et al.	2020	Survey	NA	Cross-Sectional	MBI-HSS	203	6%
Tajirian et al.	2020	EMR-Derived	Cerner	Cross-Sectional	SIBM - Mini-Z	208	45%
Tawfik et al.	2017	Organizational[3]	NA	Cross-Sectional	MBI-HSS[8]	1934	70%
Tran et al.	2019	EMR-Derived	Epic	Cross-Sectional	Mini-Z	107	56%
Rassolian et al.	2017	Survey	NA	Cross-Sectional	Mini-Z	1752	91%
Micek et al.	2020	EMR-Derived	Epic	Cross-Sectional	SIBM - Rohland	34	59%

[1] Study also included administrative data to measure work Relative Value Units (wRVUs)

[2] Manual logs were also captured to compare perceived vs actual time after hours.

[3] EMR Use was captured only at the clinic level (Use Yes or No) and correlated to burnout

[4] Intervention measured was the introduction of a Clerical Support for CPOE

[5] Intervention measured was the introduction of a scribe to the practice

[6] Intervention consisted of a clinic-focused sprint approach to optimizing EMR

[7] Authors adapted 2 single questions from the MBI-HSS that most closely correlate with Emotional Exhaustion & Depersonalization

[8] Questions related to Emotional Exhaustion and Cynicism

[9] Emotional Exhaustion sub-scale only

[10] NA=Not Applicable; NR=Not Reported; SIBM=Single-Item Burnout Measure; MBI HSS= Human Services Survey & GS=General Survey

Mini-Z survey. Many of the analyses depended on large statewide surveys such as the 2017 Rhode Island Department of Health Physician and Advance Practice Provider Health Information Technology Survey (Gardner et al., 2019), specialty group surveys (Mehta et al., 2019), or larger national surveys (T. D. Shanafelt et al., 2016) (E. R. Melnick et al., 2020). One study used manual logs to capture information about EMR use (Pozdnyakova et al., 2018), one used thematic analysis of focus group interviews (Contratto, Romp, Estrada, Agne, & Willett, 2017) and one had EMR use measured only at the organizational level (EMR yes or no) (Tawfik et al., 2017).

The studies involved where almost entirely cross-sectional (89% n=24). Of the three that measured pre- and- post-intervention, two were small single clinic studies with seven and six participants respectively. Those studies looked at clerical support for CPOE (Contratto et al., 2017) and used scribes to reduce documentation burden (Pozdnyakova et al., 2018). Despite the small sample size in those studies, both studies leveraged a version of a single-item burnout measure. Both studies used manual methods (focus group and manual logs respectively) for capturing EMR use. Only Sieja et al. (2019) did a pre-and -post intervention measurement of a larger group of physicians analyzing the impact of a clinic-focused sprint approach to optimization of EMRs for different specialty groups. They measured burnout using the Emotional Exhaustion sub-scale of the MBI as part of a broader pre-post survey strategy that highly focused on EMR use and satisfaction. Specific measures of EMR use can be found in Yan et al. (2021) except for Micek et al. (2020) who looked at time in EMR divided into four categories: weekday work hours in clinic; weekday work hours out of clinic; weekday after hours; and weekend/holiday after hours.

2.4 Discussion

The research identified in this review highlights that the definition, causes and consequences of burnout are complex and multi-faceted ranging from the individual to organizational and work related factors. In more recent years, research on the link between EMR use and burnout has grown. Evidence on individual level factors was mixed with female gender most frequently associated with more time in the EMR, and older age and higher burnout levels associated with lower perceptions of EMR usability and overall EMR satisfaction and productivity. Unfortunately, most of the literature focuses on a single specialty, with primarily descriptive evidence of variation among specialties. Because of the number of specialties and the complexity of differences between specialties, where there is evidence of variation among specialties the focus tends to be on primary care vs non-primary care specialties finding that primary care tends to have more inbox volume, higher clerical burden and the highest HIT-related stress and burnout.

This literature review did not find clear evidence that time in the EMR itself contributed to stress and burnout but did uncover abundant evidence that the perception of inadequate time for documentation did contribute to burnout. There is evidence both supporting and refuting the oft-made claim that “after-hour” time in the EMR is a major contributor to burnout. This contradiction may in part be explained by studies that take a more nuanced approach to the definition of after-hours, which found that EMR time on weeknights and administrative days did not contribute to burnout whereas time spent on weekends and holidays does. Counterintuitively, EMR efficiency tools were not found to mitigate burnout apart from the copy and paste functionality which, ironically, has been linked to note bloat, internal inconsistencies, error propagation, documentation on the wrong patient and patient safety events (Tsou et al., 2017). The fact that on-site EMR support team reduce burnout (Copley et al., 2019) may provide some evidence as to how we can mitigate the finding that negative perceptions of the EMR lead to burnout, regardless of other factors.

Research on EMR contributions to burnout is relatively new but has been rapidly increasing in the last couple of years as calls are made to move to the quadruple aim of quality that includes physician wellness as the missing quality indicator (Wallace, Lemaire, & Ghali, 2009) (O'Connor, 2015). Surveys remain the most common method of understanding how EMR use relates to burnout. This review turned up evidence of less use of the full proprietary MBI to measure burnout (33% n=9). This is remarkably lower than Rotenstein et al.'s (2018) earlier and much broader systematic review of burnout among physicians (irrespective of EMR use) which found that 85.7% of the studies included used some form of the MBI. The fact that nearly half of the studies included here use a SIBM may signal a move away from more comprehensive instruments such as the MBI. Even so SIBM's are generally not used in isolation, and instead are often supplemented with broader survey data. This overhead in administering broad surveys and difficulty in response rates for repeat measures may in part be evidenced by the limited number of interventional studies. Conversely, the lack of intervention studies may point to challenges with burnout as a dependent variable. If the goal is to reduce burnout, the research community needs a consistent and reliable way to measure the impact of interventions on burnout if they are to enact successful interventions in areas such as policy, organization change, and, most relevantly to this review, IT as have been suggested by physicians for improvement (Nguyen et al., 2021).

Finally, there is a trend away from survey instruments toward use of EMR event log data to objectively measure EMR use, a dataset that some have indicated may be a gold mine for health services research (Adler-Milstein, Adelman, Tai-Seale, Patel, & Dymek, 2020). It's worth noting that while this study shows an increase in use of event log data in the last two years, in a systematic review to define documentation burden, Moy et al. (2021) found that 80% of studies included event logging data to measure EMR use and only 20% included validation using time-motion observation. Other authors have worked to develop frameworks for leveraging these data (C. A. Sinsky et al., 2020) with recent work done to try to standardize these use measures across two major EMR vendors (Edward R. Melnick et al., 2021).

2.5 Conclusion

This literature review was designed to provide a brief survey of burnout, the impact of burnout healthcare, and how EMRs contribute to burnout. While there is no universal approach to understanding this problem, the review uncovered a trend in recent years toward use of EMR-derived data and SIBM's to measure the relationship of EMR use to burnout. The individual-level relationships to burnout as well as EMR-use relationship to burnout will serve as foundation for validating findings in the current study. While analysis on the link between EMR use and burnout is broad, I uncovered very little evidence of a learning health system approach to addressing this problem. Such an approach is needed if the industry is to respond to the call to reduce documentation burden and improve on the epidemic of physician burnout.

Chapter 3: Research Proposal

3.1 Specific AIMs

Burnout, a condition characterized by emotional exhaustion, listlessness, and an inability to cope, is demonstrably more prevalent among healthcare providers than other occupations. Burnout is fundamentally a longitudinal problem, but traditional instruments for measurement such as the Maslach Burnout Inventory (MBI) and the Professional Fulfillment Index (PFI) are time-consuming, expensive, and complex to administer, making it prohibitive to collect repeated measures to capture burnout feedback with weekly, monthly, or even yearly frequency. A Single-Item Burnout Measure (SIBM) has been previously validated to highly correlate to the emotional exhaustion dimension of the MBI, and thus can serve as a proxy measure for burnout (Rohland et al., 2004) (Dolan et al., 2015). While the causes of burnout are many faceted, including individual, workforce, and organizational factors, a large body of evidence points to Electronic Medical Records (EMRs) and subsequent clerical burden accompanied by the depersonalization of medicine due to the focus on the data at the expense of patient interaction as an underlying cause of burnout (Fred & Scheid, 2018; T. D. Shanafelt et al., 2016). To study how EMR use impacts physician burnout, it is first necessary to understand how clinicians interact with the EMR. There is widening recognition of the utility of EMR event logging data for secondary use to understand clinical workflow and EMR use patterns. Recent work proposing a non-validated framework defining EMR use in a vendor-agnostic way suggests that EMR event logging may provide a viable, scalable big data source for defining EMR use (C. A. Sinsky et al., 2020).

This proposal sought to demonstrate evidence suggesting that risk factors associated with provider burnout due to EMR use and practice, can be uncovered through judicious modeling, analysis, and assessment of EMR use metrics captured over time. The overall goal of this research was to identify EMR usage patterns that predict higher levels of physician burnout in order to design and measure just in time interventions to improve the EMR experience for individual physicians and thus mitigate burnout. To achieve this, the scope of the current proposal was to conduct a pragmatic observational trial leveraging a non-proprietary, externally validated, SIBM collected in the MU Healthcare EMR by the office of Clinician Well-Being. This SIBM was linked to relevant event logging data measuring EMR workflows to evaluate if these data collected through the EMR can be used to establish evidence of EMR workflows associated with physician burnout. To achieve this, we executed three Specific Aims (SA):

Specific Aim 1: Evaluate the feasibility and reliability of collecting repeated measures of physician burnout delivered via a novel method for no-cost, low interruption distribution of surveys through the EMR message center. We assessed provider burnout status as measured by an externally validated SIBM. The system as piloted was tested for feasibility and reliability and used to conduct studies required for Aim 3.

Specific Aim 2: Define the utility of various existing EMR workflow metrics, as captured in event logging data, through literature scan, subject-matter expert interview and data profiling. Event logging data was analyzed to define provider characterization of EMR use and identify distinguishing patterns in EMR use between specialty groups (primary care, surgery, and non-surgical medical specialties) and venue (inpatient, outpatient, and emergency department).

Specific Aim 3: Investigate the relationship between EMR use, as characterized in Aim 2, and provider burnout, as captured in Aim 1, to identify clinical workflow patterns that most highly correlate to physician burnout independent of individual physician characteristics.

Statement of Impact:

Physician burnout is a multi-faceted problem that has serious human consequences for physicians in addition to potential adverse patient safety impacts and financial impacts on an already overburdened healthcare system. If Healthcare Information Technology is to reduce the clerical burden on physicians, we must first understand what aspects of the system are most highly correlated to physician burnout. This understanding is foundational to future work to predict burnout based on more granular EMR use patterns and to subsequently design and deploy interventions with measurable reduction of physician burnout.

Research Strategy

3.2.1 Significance

"Burnout" is a condition characterized by emotional exhaustion, listlessness, and an inability to cope (Informed Health, 2020). While burnout is found across all occupations, this condition is significantly more prevalent among healthcare workers than in the general population, affecting as many as half of U.S. nurses and physicians (Tait D. Shanafelt et al., 2012), with an estimated impact of approximately $4.6 billion in annual costs to health systems related to physician turnover and reduced clinical productivity (Han et al., 2019). In fact, evidence also shows that the gap between physician risk of burnout and decline in work-life balance relative to the general population has worsened in recent years, even after adjusting for factors such as age, relationship status and hours worked per week (T. D. Shanafelt et al., 2015). Systematic review reveals the risk to physicians and onset of burnout has a complex etiology involving a) individual factors such as age, gender, relationship status, sleep deprivation and self-coping; b) work factors such as specialty, long hours, job position, interpersonal demand, job security and limited resources; and c) organizational factors such as negative leadership, workload expectations, and inadequate rewards (Azam et al., 2017; Patel et al., 2018). Electronic Medical Records and subsequent clerical burden, as well as the depersonalization of medicine due to a focus on complete, accurate and timely data input at the expense of patient interaction, has repeatedly been identified as an underlying cause of burnout (Fred & Scheid, 2018; T. D. Shanafelt et al., 2016).

Burnout is a fundamentally longitudinal phenomenon, meaning that as the underlying individual, work and organizational factors evolve, feelings of burnout will evolve (Grumbach et al., 2019; Rinne et al., 2020). Consequently, identification of risk factors for burnout and measurement of the impact of improvement efforts necessarily requires data collection that can be repeated in short duration such as weeks, months, or years.

The current standard for measuring burnout is the Maslach Burnout Inventory (MBI) which is derived from a 22-item survey (Maslach et al., 1986; Maslach et al., 2016). This instrument is well validated, has decades of use, and measures burnout on three natural dimensions: emotional exhaustion, depersonalization, and personal accomplishment. However, the MBI is proprietary and carries licensing fees. In 2018 a study out of Stanford proposed the Professional Fulfillment Index (PFI) as a free alternative to the MBI that captures professional fulfillment in addition to burnout mechanisms that are not captured by the MBI (Trockel et al., 2018). However, both the MBI and the PFI are complicated and time-consuming to administer, making practical use for collecting timely longitudinal feedback on physician well-being prohibitive. To address the complexity, time and cost issues, Single Question Burnout Surveys (SQBS) were created and validated against the larger surveys providing an alternative validated burnout measure with far less cost and complexity and lower burden on the respondents. **One alternative, a Single Item Burnout Measure (SIBM), has been demonstrated to validate against the emotional exhaustion dimension of the MBI, making it a viable, low-cost proxy for the MBI in evaluating the emotional exhaustion component of burnout** (Rohland et al., 2004; Dolan et al., 2015).

The 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act offered financial incentives for providers who exercised Meaningful Use of Certified Electronic Health Record (CEHRT) technologies. Since this led to almost universal adoption of CEHRT (Mennemeyer, Menachemi, Rahurkar, & Ford, 2016) including patient Protected Health Information (PHI), event logging data was originally mandated by the Security Rule in the Health Information Portability and Accountability Act (Scholl, 2008). Soon after implementation, the use of event logging data acquired within the EMR systems became an important means to study clinical workflows and EMR practice (C. A. Sinsky et al., 2020). Event logging data typically includes EMR use metrics reflecting provider interaction with the patient’s information as measured by their interaction with the EMR (e.g., counts of events, amount of time per event, and average over unit of time). Researchers have increasingly identified observational use cases to leverage this data to measure physician activity (Rule, Chiang, & Hribar, 2020).

As part of the requirements for the 21st-Century Cures Act, in February 2020 the U.S. Department of Health and Human Services released a report examining the administrative burden of Health IT and EMRs. They called for reduced effort and time required to record information, reduction of effort and time to meet regulatory requirements, and improved ease of use for EMRs (U.S. Department of Health & Human Services, 2020). A recent scoping review of studies designed to reduce the impact of EMR documentation burden found that 80% of the articles used event logging data whereas only 20% included some validation to a known previous measurement or time-motion study. They also found that most of the studies tied back to some measure of clinician satisfaction or burnout (Moy et al., 2021) . However, standardized measurements across systems have not been defined to support comparative effectiveness evaluation of documentation improvement or EMR usability efforts (C. A. Sinsky et al., 2020). This proposal is the first to propose the design and validation of a system using an in-workflow distribution of a burnout measure. We further propose to demonstrate that this burnout measure, coupled with use of event logging which captures EMR activity, can be used to support a Learning Health System (LHS) characterized by “continual improvement and innovation” with “new knowledge captured as an integral by-product of the delivery experience” (Horwitz, Kuznetsova, & Jones, 2019; Friedman et al., 2014; McGinnis, Stuckhardt, Saunders, & Smith, 2013).

3.3 Innovation

This study represents the first work to design and validate a scalable platform for understanding 1) how the EMR impacts physician burnout, 2) how EMR usage can be used to predict future burnout, and 3) providing a way to target and measure the impact of improvement interventions. The body of research examining physician burnout, the underlying causes, and the downstream impacts provides evidence that this is a growing, high-impact problem in an already overburdened healthcare system. This is the first attempt to test the feasibility and reliability of collecting a single item measure of physician burnout directly in the clinical workflow. It is also the first work to develop a system for tying that burnout back to clinical workflow via event logging data to demonstrate system-wide correlation of EMR workflow to burnout. The system is designed in a way that will allow for iterative deployment and measurement of future targeted improvement efforts within an LHS framework.

Because the scope of this work is contained within a single AMC, MU Health, any findings about the relationship between clinical workflow and burnout are not generalizable. Rather the scope of this work validates the feasibility of using these tools as a platform for continuous improvements within a LHS in a way that will enable future innovation. While EMR improvement efforts are continuous within the hospital, future innovation will lie in our ability to identify targeted improvements and evaluate the impact of those improvements. This work paves the way for future innovation including: the development of advanced modelling of raw event logging data and use of machine learning to develop novel predictors of burnout; scaling of measurement and improvement efforts across sites to gain generalizable knowledge with larger populations;

and collaboration to standardize the way in which event logging data is captured and prepared in a way that scales is generalizable across EMR vendors.

Table 3. 1: Helpful acronyms and their corresponding terms

Key Acronyms	
OCW	Office of Clinical Well-Being
EMR	Electronic Medical Record
SIBM	Single-Item Burnout Measure
LHS	Learning Health System
HITECH	Health Information Technology for Economic and Clinical Health
CEHRT	Certified EHR Technology

3.4 Approach:

3.4.1 Background

To accomplish the Aims of this proposal, we have partnered with a Quality Improvement team at MU Health led by the Chief Clinical Wellness Officer. MU Health is a 650-physician Academic Medical Center located in Columbia, MO that serves a 25-county catchment area throughout Mid-Missouri. Buoyed by a public/private partnership with Cerner Corporation called the Tiger Institute, MU Health has a long history of experience in quality improvement around the clinician EMR experience. However, evidence of the impact of these improvements is generally anecdotal or process-based. MU Health has a strategic focus on improving the EMR clinician experience and has administered an internal, anonymous physician and nurse EMR experience survey for each year from 2015-2017. In 2018 MU Health moved to the KLAS ARCH Collaborative (KLAS, 2020) to gather provider feedback on the EMR experience in a way that could be benchmarked against other organizations. However, this survey costs money and does not purport to measure the impact of any specific intervention, including burnout. Closing this gap requires a validated instrument for capturing repeated measures of physician burnout on a weekly, monthly, or yearly basis in a way that is practical and cost-effective.

Earlier this year, the Office of Clinician Well-Being (OCW) partnered with an on-site application development team to implement a Single-Item Burnout Measure (SIBM) in the EMR as part of overall improvement efforts to address clinician wellness at MU Health. During idea formulation, the (OCW) engaged this research team to formulate a research design with the goal of identifying innovative new ways to reduce EMR causes of clinician burnout.

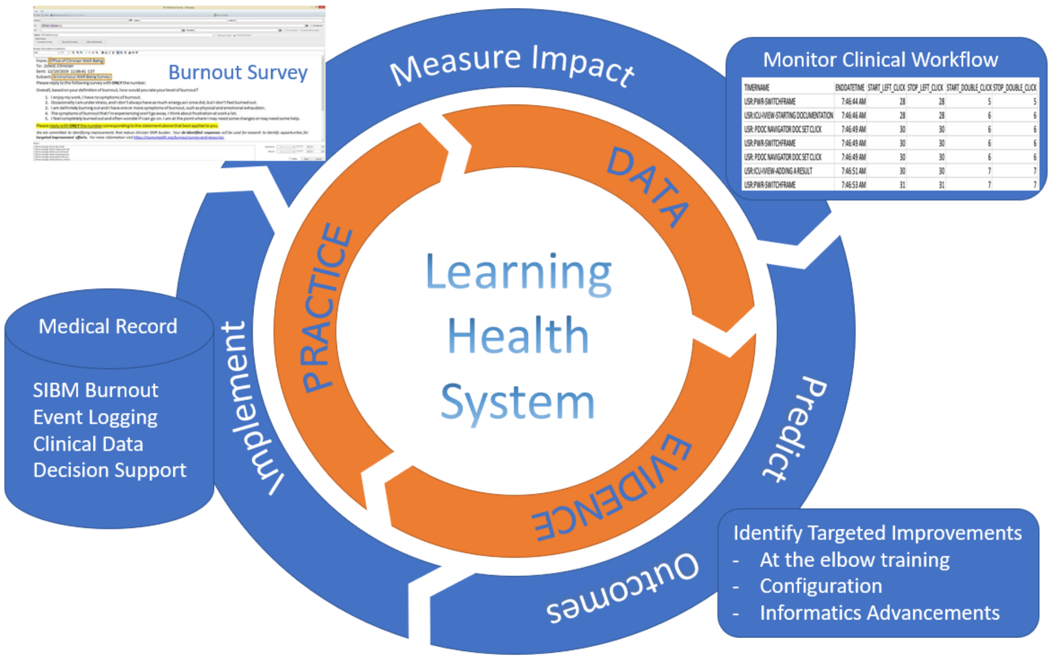
3.4.2 Conceptual Framework

This proposal approaches the problem of EMR clinician burnout within the conceptual model of a LHS (Figure 3.1), a vision set by the Institute of Medicine (IOM) for a healthcare system that engages in continuous learning through an emphasis on better use of data (Institute of Medicine, 2020). To do so requires a validated process for continuous identification of targeted improvement opportunities, implementation of improvement into the clinical workflow, and measurement of impact.

3.4.3 Scientifically Relevant Preliminary Studies:

This proposal uses two datasets: aggregated event logging data and self-assessed burnout as measured through a previously validated SIBM administered through the EMR. Three possible mechanisms were identified to meet the OCW goal to capture feedback simply and repeatedly on clinician well-being in the EMR: interruptive decision support, non-interruptive decision support (nudges), or the physician message center. Interruptive decision support was expected to be a dissatisfier and therefore might skew results. Unfortunately, technology limitations did not allow for non-interruptive decision

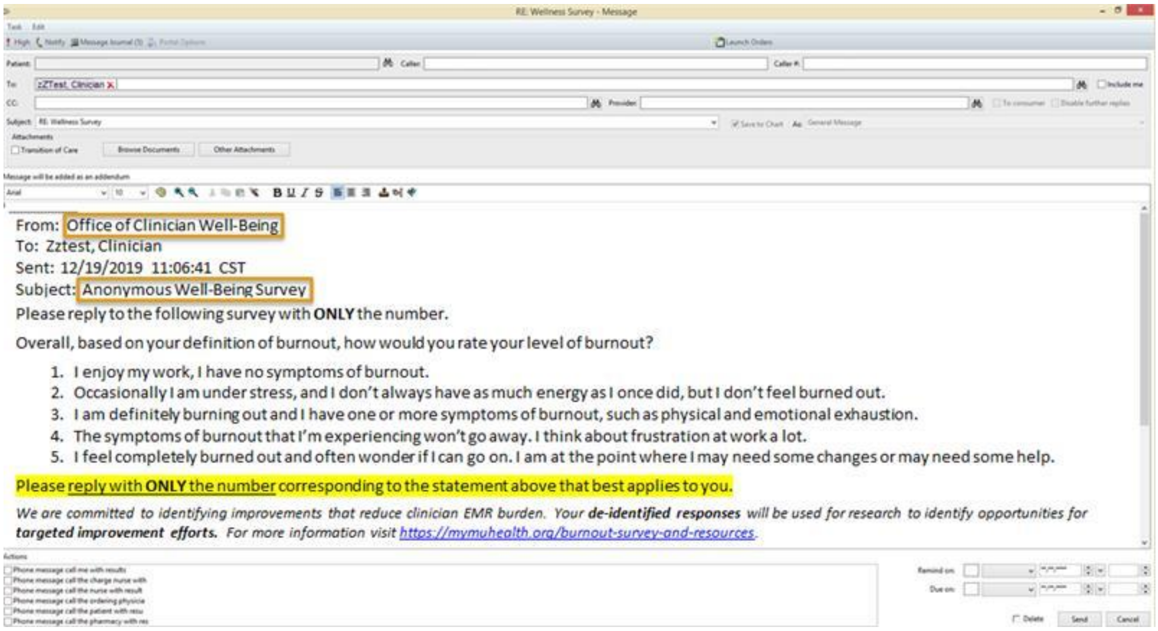
Figure 3. 1: Conceptual model for research design



support, so the EMR message center was selected for delivering the survey.

A successful pilot was conducted on a convenience sample of 15 physicians and led by the Medical Director for decision support. A follow-up survey completed by 13 of the 15 pilot participants focused on usability, utility, and privacy. Figure 3.2 shows the pilot survey questions and responses. Users scored usability as an average of 4.9 on the Single Ease question. This single-item usability question asks: “Overall how difficult or easy was this task to complete.” In general users indicated that the survey is an effective and accurate reflection of their overall feelings of burnout with one user indicating in comments that they intentionally gave an inaccurate response since this was a pilot for feedback. The greatest concerns raised in both the codified and open-ended responses concerned privacy. Based on the feedback, privacy and de-identification was emphasized

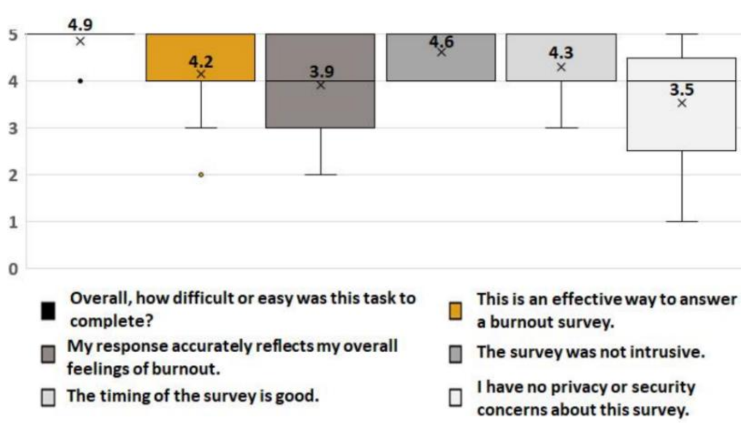
Figure 3. 2: SIBM as distributed through the EMR Message Center.



in the communications from leadership, linked documentation, and the survey itself. Additionally, the QI team identified frequent purge processes to ensure that data did not persist unnecessarily in the EMR system where it would be visible to others.

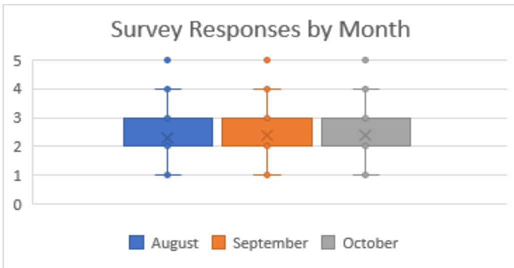
In August through October of 2020, The SIBM as designed was administered once a month for three months to MU Health attending physicians across all specialties and venues via the EMR message center inbox. To facilitate future research design and control for bias related to day of the week and time of day each month physicians were randomized into day of week and time of day cohorts. Summary results in Figure 3.4 show an average response rate of 38% with the highest

Figure 3. 3: Pilot survey responses



responses in the first month and decreasing response rates in the second and third surveys. Survey responses and the corresponding distribution appears to remain relatively stable month over month.

Figure 3. 4: SIBM preliminary response rates and average scores collected in repeated survey of all attending physicians at an AMC over a period of three months.



Month	Valid Responses	Surveys Sent	Response Rate
Aug	286	649	44%
Sep	247	627	39%
Oct	197	623	32%

In addition to the SIBM, MU Healthcare's EMR captures event logging data as mandated by HIPAA. Originally designed to replace paper records and store demographics and clinical documentation, current day EMRs are dominated by large-scale EMR vendors. Partially buoyed by incentives for government funding, these EMRs support workflows across both venues (inpatient, emergency, and ambulatory) and specialties. The event logging systems have been extended to allow us to capture information about time spent on various activities, levels of adoptions, and workflow patterns. In the MU Health EMR, this is achieved through the Response Time Measurement System (RTMS) timers built into the software to monitor activity in the EMR. It consists of timers that activate when the user performs either more than three mouse clicks or 15 keystrokes or 1700 mouse miles per minute (movement the cursor

Table 3. 2: Prescribing providers captured in Cerner Advance event logging data

Position	Total Providers
Attending Provider	615
Fellow	15
Resident	495
Physician's Assistant/Advanced Practice	152

travels across a monitor) (Aziz, Talhelm, Keefer, & Krawiec, 2019). These event logging data are pre-processed into approximately 400 measures by the EMR vendor and made available, primarily for quality improvement, through the LightsOn Network (https://www.cerner.com/solutions/lights-on-network) and Cerner Advance (https://advance.cerner.com/). In total, these systems capture 400 aggregate variables of provider workflow and include information on 1277 prescribing providers (Table 3.2). Sample Aggregate data were collected for the month of February and evaluated for distribution, quality, and availability. Based on this assessment, literature review, and subject matter expert interview, 16 variables were identified that we proposed are representative of the EMR workflow and experience (Table 3.3).

Table 3. 3: Proposed high-level EMR metrics (See Appendix A for data dictionary)

Measure	Description
Total time in EMR	Total hours spent in the EMR during and outside of the clinical visit
Chart Review Time	Hours spent reviewing the flowsheets, summary pages, and clinical notes
Documentation Time	Hours spent creating the physician clinical note
Prescription Time	Hours spent on orders
Message Center Time	Hours spent in the EMR inbox messaging members of the care team, external providers, or the patient
Health Maintenance Time	Hours spent managing chronic conditions which may extend beyond just the acute reason for visit
Patient Discovery	Hours spent managing patient lists, tracking list and locating patients
Work Outside of Work	Measure of time spent in the EMR outside of core work hours
Teamwork for Orders	Percentage of orders signed by the originating physician. Designed to capture order that originate from residents or protocol
Electronic Documentation Adoption	Percentage of notes that originated in the EMR and were electronically signed by the physician
EMR Performance	Average transaction response time across all EMR activities
EMR Stability	Percent of time in the EMR in which the physician experiences an application crash
Decision Support Interaction	
Total Alerts Fired	Total number of decision support alerts experienced by physician
Percent Overriden	Percent of overridable alerts that were overridden by the physician
Mobility	Captures the extent to which documentation happens on a mobile (iPad, Surface, smart phone)
Notes	Total notes signed in mobile application
Orders	Total Orders signed in mobile application
Charts Opened	Total charts opened in mobile application

3.4.4: Specific Aims

<u>Overview of Aim 1</u>:

Evaluate the feasibility and reliability of collecting repeated measures of physician burnout delivered via a novel method for no-cost, low interruption distribution of surveys through the EMR message center. Provider burnout status was measured using an externally validated Single Item Burnout Measure. The system as piloted was then tested in a larger sample for feasibility and reliability and used to conduct studies required for Aim 3.

<u>Aim 1.1: Evaluate the likelihood of physician participation in a time and day randomized, EMR-administered SIBM to identify patterns of bias and confounding factors</u>:

The utility of administering the SIBM through the EMR is dependent on non-biased survey responses. To test this, we examined the impact of individual-level factors (gender, age cohort and specialty group) as well as environmental factors (day of week and time of survey administration, day of week and time of response, and venue) on physician likelihood of response as well as on response rates (0, 1, 2, or 3 total responses) using univariate statistical techniques across the entire surveyed population (n=620). Table 3.4 lists out all relevant survey and physician attributes. Chi-square with Mann-Whitney U test tests (Wilcoxon rank sum) for two-level independent variables and Kruskal-Wallis for greater than two-level were used for statistical tests. Non-parametric tests were used because the dependent variable, response rates, were not normally distributed. Significant results are evidence that the independent variable (day and time of distribution) independently impacted the likelihood a physician will respond to the

Table 3. 4: Individual physician and survey metadata

Survey Attributes	
Day of Week Administered	**Randomly assigned day of week that survey was sent**
Time Administered	**Each physician was randomly assigned to either early morning (7:00 am), mid-morning (11:00 am), or afternoon (3:00 pm)**
Day of week Responded	**Day that physician responded to the survey**
Time Responded	**Time of day that physician responded to the survey**
Response Value	**SIBM burnout response on a scale of 1-5**
Physician Attributes	
Age	**Physician categorical age at time of survey (< 40, 40-55, > 55)**
Gender	**Physician gender**
Specialty Group	**Primary care, surgical or on-surgical medical specialty**
Total Responses	**Number of times the physician responded (0,1,2, or 3)**
Average Response	**Average SIBM burnout score across the three month survey period**
% Ambulatory	% of time in an ambulatory setting as measured by time in EMR
% Emergency	% of time in an emergency setting as measured by time in EMR
% Inpatient	% of time in an inpatient setting as measured by time in EMR

survey or burnout. This may advise the best day and time for administering the survey to maximize response rates in the future.

<u>Aim 1.2: Test the reliability of an SIBM administered through the EMR</u>:

Because our study design relied on previous research externally validating the SIBM as a valid measure of physician burnout, this study design focused on the internal reliability of the measure. This test – re-test method relies on the assumption that any individual physician experiences approximately the same level of burnout from one month to the next. To test the internal reliability, we calculated the intraclass correlation Coefficient (ICC) on the ordinal response variable measure of burnout (1-5) between month 1 and 2 and month 2 and 3 responses for individual respondents. ICC can be calculated for responses when tested over more than two time periods. ICC can be calculated as (Koo, 2016):

$$ICC = \frac{MS_{Subject} - MS_{Error}}{MS_{Subject} + (k-1)MS_{Error}}$$

Where k = the number of time periods, $MS_{subject}$ = the between-subjects mean square, and the MS_{error} = the mean square due to error after fitting a repeated measures ANOVA. Sample size was limited to only those physicians who responded more than once and in consecutive months. In other words, month 1 was compared to month 2 and month 2 to month 3 where responses are received for a single physician in consecutive months. Sample sizes were adequate for ICC tests. The ICC results were further confirmed through a wilcoxon signed rank test.

Aim 1.3: Test the impact of individual and environmental-level factors on burnout responses

While the SIBM has been validated as a surrogate for Emotional Exhaustion (what we are terming as "burnout" for the purposes of this study), we do not know of previous work to look at the impact of survey administration on feelings of burnout. Because time and day of survey administration are randomized, the expectation was that they would not be correlated feelings of burnout (response variables 1-5). Because each physician was randomized each month into a day of week and time of day cohort, we proposed to test using each response as an independent observation. Two accomplish this, we conducted chi-square tests with Kruskal-Kallis to test for significant differences between the independent variables. This was reinforced using a one-way ANOVA to test if the means are equal across the randomized variables (day of week and time of day). The expectation was that there is no relationship between day and time of administration and burnout.

Aim 1 Anticipated Results and Alternative Strategy:

We expected to accept the null of no relationship between individual characteristics and response rates but suspected we would find some biases. Response rates across the three-month study period were high enough that we did not anticipate any sample size issues, but any sub-group sample sizes < 6 would not be included in final analysis to protect the identity of the physicians. We expected that any identified patterns of bias in this will advise Aim 3. For example, if there was clear bias or low response rates in an identified sub-group, Aim 3 needed to be adjusted for the relevant sub-population (e.g., just primary care). Because they were randomized, date and time of administration should have no relationship to the burnout response itself. In the event we find a relationship, Aim 3 would have needed to be adjusted for this unexplained confounder.

Overview of Aim 2:

Define the utility of existing EMR workflow metrics, as captured in event logging data, through literature scan, subject-matter expert interview and data profiling. Event logging data was analyzed to define provider characterization of EMR use and identify distinguishing patterns in EMR use between specialty groups (primary care, surgery, and non-surgical medical specialties) and venue (inpatient, outpatient, and emergency department).

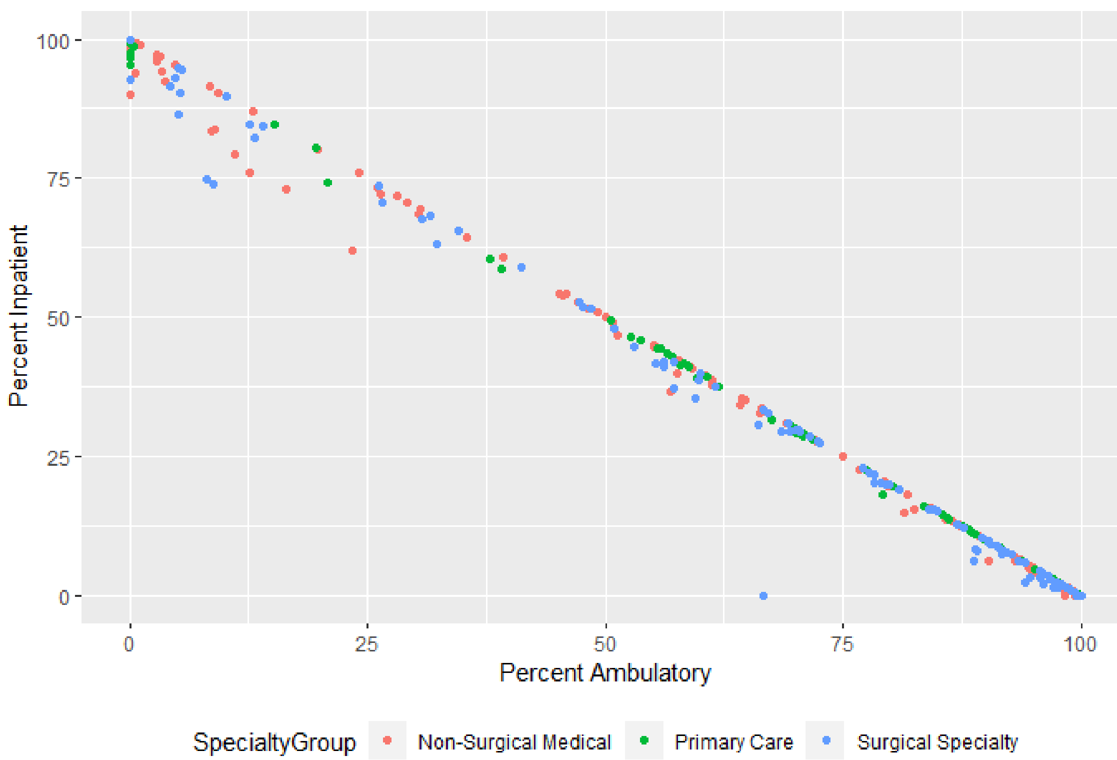
Aim 2.1: Characterize patterns in EMR usage factors that are significant in distinguishing inpatient vs ambulatory vs emergency department venues:

Preliminary data shows that physicians cannot easily be categorized by the venue in which they practice (Figure 3.5). Instead, most physicians work with patients in multiple venues. To understand how clinical workflow patterns varied between venues, we used ICC to test if there is stronger in-provider correlation across the variables of interest or if there is stronger correlation within venues across providers. Since our analysis in Aim 3 is at the provider level, understanding these underlying patterns was necessary to make decisions about how to operationalize the data for Aim 3. We hypothesized that an individual physician's EMR workflows remained consistent across venues, but that as a group there is definable variance between inpatient, ambulatory, and emergency department EMR usage.

Aim 2.2: Characterize EMR usage factors that are significant in distinguishing between primary care, surgical specialties, and non-surgical medical specialties.

In this aim we used factor analysis to identify latent EMR usage patterns to test the hypothesis that EMR use varies across venue and specialty and define those EMR

Figure 3. 5: Preliminary data illustrating percent of time spent in ambulatory vs inpatient based on time spent documenting on ambulatory or inpatient encounters respectively



patterns that characterize each specialty. In part to address small sample sizes in certain specialties, we aggregated specialties into meaningful sub-groups, specifically: primary care defined as pediatrics, general internal medicine, and family community medicine;

surgical specialties; and all other non-surgical medical specialties. Ancillary specialties such as pathology and radiology were excluded due to low or no use of the EMR at MU Health. We first identified which of the relevant features above related to specialty group using intra-class correlation (defined above) to identify those EMR workflows that distinguish between specialty groups. Next, we used factor analysis to identify latent patterns of clinical workflow. Logistic regression was used to identify if the latent classes significantly correlated to venue and specialty.

Aim 2: Anticipated Results and Alternative Strategy:

We anticipated that physician EMR workflows would remain constant for an individual physician across venues, suggesting that all EMR metrics can be aggregated at the physician level for Aim 3 and suspected that if this is not the case, Aim 3 needed to be adjusted to look at relationships within each venue separately. Similarly, variations in workflow identified across specialty groups advised the analysis strategy for Aim 3 since variations may indicate that signals not apparent at the overall level may prove to be significant at the sub-group level. It may be that alerts and overrides are a common characteristic, for example, in medical specialties or inpatient venues, and less so in surgical or ambulatory specialties. Such a finding would suggest that separate analysis is needed at the subgroup level.

Overview of Aim 3

Investigate the relationship between EMR use, as characterized in Aim 2, and provider burnout, as captured in Aim 1, to identify those clinical workflow patterns that most highly correlate to physician burnout independent of individual physician characteristics.

Aim 3.1: Evaluate correlation between EMR metrics and physician burnout

The EMR SIBM was sent to identified providers three times, once a month, over the course of three months. Day of the week and time of day of survey administration were randomly assigned to respondents. We anticipated that different cohorts would respond each month. These response data were compared to the physicians' EMR usage patterns each month. Relevant individual and survey meta data are detailed in Table 3.4.

These data were combined with the relevant EMR workflow metrics identified in AIM 1. Univariate regression modelling was used on each physician month-response pair to identify those variables that most highly correlated with physician burnout. We expected to have to account for the repeated measures either though a random effect for physician, by selecting the first response month for each physician, or, if measures are stable over time as measured by ICC, by averaging the months per physician. Change in univariate regression relationships were also tested in multivariate models while controlling for individual-level characteristics such as gender, age, and specialty.

Aim 3 Anticipated Results and Alternative Strategies

Based on the results of Aim 1 and 2, we anticipated that adjustments would need to be made to the analytic approach. Specifically, we anticipated two areas that would require change to the planned approach. First, the dependent variable (DV) represented by the SIBM may not be adequately sensitive. Further, the DV is skewed to the right and required that we create a binary variable of burnout. Second, the use of aggregate event logging data can potentially attenuate signals that may be found in the more detailed underlying data. In both cases, the likely result was that no correlation would be found between EMR use and burnout when controlling for individual factors, despite a strong body of evidence that there is a relationship. For example, based on existing literature, we expected to find a strong correlation between work done outside the hours of 6:00 am and 6:00 pm and burnout scores but expect this signal will only be found in the ambulatory physicians because of shift work. Any non-results dictate future direction in how we measure physician wellness and how we operationalize event logging data to predict need for and design interventions for improvement.

E. Ethical Considerations

E.1 Human Subjects Research

Because physician burnout addresses mental health, special care was taken to ensure the privacy of the participants included in the retrospective study. For the purposes of preventing any potential re-identification of any individual physician, attributes identifying any group with size < 6 was suppressed from the final de-identified datasets provided to researchers by the honest broker. Because the data for this study is entirely de-identified, this work was determined to not constitute Human Subjects Research (HSR) by the MU Institutional Review board.

E.2 Data Safety Monitoring Plan:

Data collected through the course of quality improvement activities at MU Healthcare is being harvested from the MUHC EMR and stored at the physician level in a custom table accessible only on the backend of the EMR. All messages in the EMR (received and sent) were regularly purged from the application and database. The MUHC Living lab team served as an honest broker to combine data sources at the physician level and removed any identifiable information before providing these data to the research team. Any textual responses were manually scrubbed for identifiable information and provided to investigator for qualitative analysis. The final dataset was fully de-identified for research and as such can be shared broadly along with the data dictionary to allow for external validation of the work contained in this proposal.

Chapter 4: Single-item burnout measure in the clinical workflow

4.1 Abstract

While documentation burden is driven by several valid regulatory, quality, and patient safety factors, it continues to contribute to physician burnout with high impact on physician well-being. As the research community turns their focus to continuous improvement to combat this trend, a reliable method to capture repeated measures of physician burnout is necessary to measure the impact of improvement efforts. In late 2020, MU Healthcare deployed an externally validated Single-Item Burnout Measure (SIBM) directly into the clinical workflow to measure the reliability of capturing repeated measures of burnout to aid pre-post intervention studies. Chi-square for independence of nominal variables and Mann-Whitney U (Wilcoxon rank sum) Test for p-values of variables with two levels and Kruskal-Wallis for p-values with > 2 levels were used to test the impact of survey administration (day of week, time of day, and month) and individual characteristics (age, gender, and specialty) on both response rates and burnout scores. Gender (p=.0098) and specialty (p<.001) were related to response rates with 99% confidence and gender (p=.002), age (p=.001) and specialty (p=.012) were significant for burnout score. No survey administration variables were significantly related to burnout. Intra-Class Correlation (ICC=.871) and Wilcoxon signed rank tests (p=.307) supported good test-retest reliability. An SIBM deployed in the workflow provides a feasible and reliable way to capture repeated measures of burnout needed for measurable continuous improvement.

4.1 Introduction

Burnout is defined as a condition characterized by listlessness and an inability to cope. Shanafelt et al (2012) found that the condition is significantly more prevalent among US nurses and physicians, affecting as many as half (Tait D. Shanafelt et al., 2012). One systematic review of attending physicians only found that as many as 85% of physicians are impacted by some dimension of burnout (Rotenstein et al., 2018). Burnout has been demonstrated to impact individual job satisfaction, sick leave, and turnover leading to lower clinical productivity, lower patient satisfaction and adverse clinical outcomes. One much-cited study estimates the cost of burnout to the overall health system to be $4.6 billion annually (Han et al., 2019).

The gold standard for measuring burnout is the Maslach Burnout Inventory (Maslach et al., 1986). This proprietary instrument measures burnout on three dimensions: emotional exhaustion, depersonalization, and personal accomplishment. Free alternatives such as the professional fulfillment index (PFI) capture additional dimensions not captured by the MBI such as professional fulfillment, work exhaustion, and interpersonal disengagement (Trockel et al., 2018). Other groups have sought to demonstrate single-question measures such as the Single-Item Burnout Measure (SIBM) and the Mini-Z's Single-Question Burnout Score (SQBS) as valid and simpler alternatives to full surveys (Dolan et al., 2015; Olson, Sinsky, et al., 2019; Rohland et al., 2004). One systematic review found that in 182 studies in 45 countries that there were 142 different ways to measure burnout, many of which are costly and time-consuming to administer making them less useful for measuring changes in clinician burnout over time (Rotenstein et al., 2018).

The May 2021 special issue of JAMIA focused on many facets of clinician burnout and potential interventions including Health Information Technology (IT) policy, Health IT interventions, and organization-directed improvements. Even with the focus on improving clinician burnout, these works collectively highlighted the challenge of capturing individual level data on clinician burnout over time to measure whether interventions were having the intended effect on clinician burnout. To close that gap, we deployed an externally validated SIBM within the Electronic Medical Record (EMR) workflow to demonstrate the feasibility of this method to collect repeated measures of individual clinician burnout.

4.1.1 Ethical considerations

The current study was approved by the institutional review board of the University of Missouri as Quality Improvement. The study team instituted a de-identification protocol to mask any sub-groups with an n < 6 to ensure protection of the clinicians' identities.

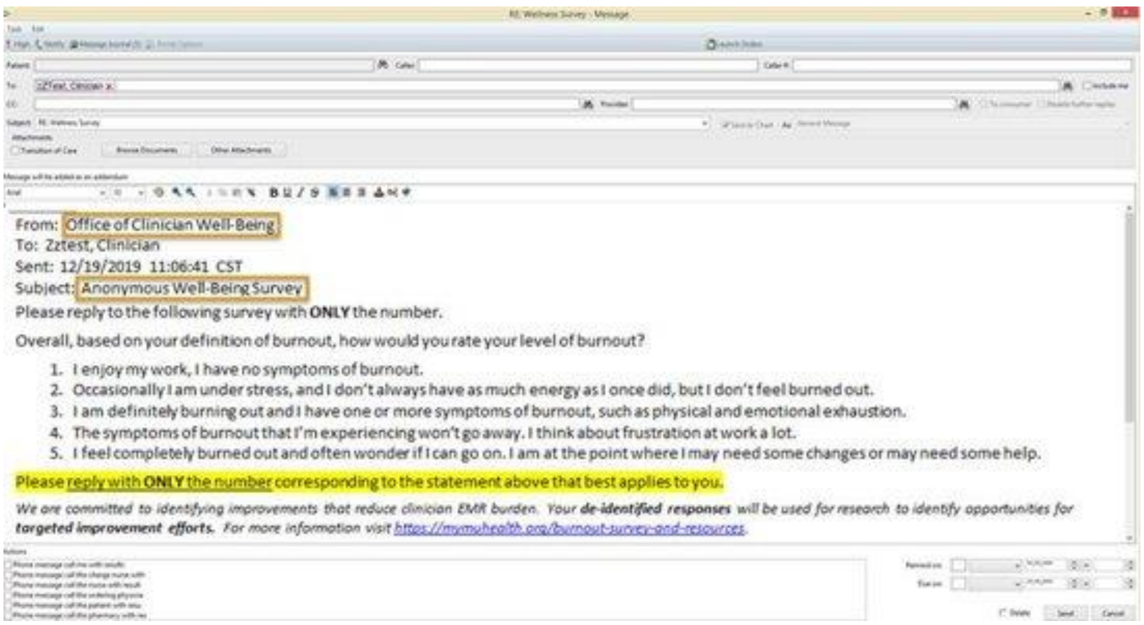
4.2 Materials and Methods

The study team set out to test the feasibility and reliability of administering a SIBM via a novel method for no-cost, low interruption distribution of surveys through the EMR Messaging Tool. The study team consisted of leaders from the MU Health Office of Clinical Well-Being, the Medical Director for decision support, Industrial Engineering and Health Informatics experts, and a local software engineering team.

4.2.1 SIBM Instrument and pilot

For this study, we selected an existing SIBM that has been previously externally validated to the Emotional Exhaustion dimension of the MBI (Rohland et al., 2004). The survey asks “Overall, based on your definition of burnout, how would you rate your level of burnout”. Figure 4.1 shows the five-point response scale ranging from no symptoms of burnout to “I feel completely burned out and often wonder if I can go on”.

Figure 4. 1: The SIBM delivered through the EMR Message Center



4.2.2 Survey Administration

The team early on discarded the possibility of using an interruptive decision support alert on the premise that the alert itself may bias the clinician’s responses. The technical team determined that there was not an existing way to deliver a non-interruptive decision support alert or a “nudge”; therefore, the study team chose the EMR e-mail-like messaging tool, “message center”, to administer the survey and collect responses. Prior to deployment of the survey, a convenience survey of 15 physicians was conducted by the Medical Director for decision support. Overall, the pilot group scored the survey very high (4.9 out of 5) on usability as measured by the Single-Ease usability question, which has been demonstrated to be the quickest and easiest to administer (Sauro & Dumas, 2009) and more reliable in small sample sizes (Tedesco & Tullis, 2006). Respondents also indicated that the survey was an effective (4.2 out of 5) and non-intrusive (4.3 out of 5) way to answer a burnout survey. The feedback from the pilot did raise concerns about both the timing of the study and privacy and security around collecting burnout responses. To address those concerns, a communication plan was developed with the hospital and school of medicine leadership. These communications explained the goal of reducing burnout and provided more information on how confidentiality would be maintained. The survey itself was also amended to emphasize the anonymous and de-identified nature of the study results. The technology team developed a process to automatically pull the results from the EMR each night. These results were de-identified by an honest broker according to our de-identification protocol prior to being made available to the study team. There was some concern over how well the SIBM reflects the level of burnout (3.9 out of 5), but no changes were made to the survey since we are relying on previous evidence of the validity of the chosen measure.

To address the potential impact of the time of administration on the survey, clinicians were randomized into day- of- week and time- of- day cohorts. The survey administration was limited to physician providers because use of the message center is less common among nursing staff,   The survey was administered to all attending physicians once a month for three months from August to October of 2020 at a single midwest academic medical center. Individual characteristics such as age, gender, and specialty were linked to the data set. To further ensure that we protect privacy, specialties were grouped into primary care, emergency, surgical and non-surgical medical specialties (See Appendix D of supplementary materials). Table 4.1 shows the survey and individual physician attributes that were captured. Chi-Square tests for nominal variables with Mann-Whitney U test (for two categories) and Kruskal-Wallis (for > 2 categories) were used for p-values. ICC and Wilcoxon signed rank tests were used to examine test-retest reliability from month to month. All statistical tests were run using R version 4.1.0 with IRR package for ICC and Stats for all other statistical tests (Gamer, 2019; R Core Team, 2021).

Table 4. 1: Survey and provider characteristics

Survey Attributes	
Day of Week Administered	Randomly assigned day of week that survey was sent
Time Administered	Each physician was randomly assigned to either early morning (7:00 am), mid-morning (11:00 am), or afternoon (3:00 pm)
Day of week Responded	Day that physician responded to the survey
Month of Response	August, September, October, 2020 (Month 1, 2, and 3)
Total Responses	Number of times the physician responded (0,1,2, or 3)
Response Value	SIBM burnout response on a scale of 1-5
Physician Attributes	
Age	Physician categorical age at time of survey
Gender	Physician gender
Specialty Group	Defined as either primary care, Surgical, non-surgical medical, or emergency

4.3 Results

4.3.1 Impact of Population Characteristics and Survey Administration on Response rates

Table 4.2 shows the physicians surveyed by specialty and demographics. In total, 535 physicians were surveyed with 499 receiving the survey all three months. The survey group was skewed somewhat toward males and the non-surgical medical specialties – in part likely due to masking of individual attributes of smaller sub-groups. Females had significantly higher response rates than males (p=.0098) while primary care had significantly higher response rates than other specialties (p<.001). (Table 4.3). Randomized day of week and time of day did not significantly impact response rates. However, there was a significant drop-off in response rates over time from month one to month three (Table 4.4).

Table 4. 2: Total physicians surveyed by specialty and demographics

Category	Total Surveys Sent 3	2	1	Total (% of Total)
Total Physicians	499	22	14	535
Gender				
Female	180	6	4	190 (36%)
Male	272	4	2	278 (52%)
NA[1]	47	12	8	67 (13%)
Age Group				
25-39	146	4	1	151 (28%)
40-54	217	4	4	225 (42%)
55+	113	3	2	118 (22%)
NA[1]	23	11	7	41 (8%)
Specialty Group				
Primary Care	127	1	1	129 (24%)
Non-Surgical Medical	210	8	5	223 (42%)
Surgical	124	10	6	140 (26%)
Emergency	38	3	2	43 (8%)

[1]For any sub-group less than 6, demographics were masked to protect identities of the physicians

Table 4. 3: Response rates by specialty and demographics

Category	Total Responses[1] 0	1	2	3	p-value[3]
Gender[2]					**0.0098****
Female	44 (24%)	27 (15%)	35 (19%)	74 (41%)	
Male	92 (34%)	51 (19%)	42 (15%)	87 (32%)	
Age Group[2]					**0.5407**
25-39	39 (27%)	28 (19%)	33 (23%)	46 (32%)	
40-54	68 (31%)	37 (17%)	32 (15%)	80 (37%)	
55+	33 (29%)	13 (12%)	19 (17%)	48 (42%)	
Specialty Group					**< .001****
Primary Care	23 (18%)	13 (10%)	22 (17%)	69 (54%)	
Non-Surgical Medical	71 (34%)	40 (19%)	33 (16%)	66 (31%)	
Surgical	43 (35%)	19 (15%)	28 (23%)	34 (27%)	
Emergency	15 (39%)	10 (26%)	4 (11%)	9 (24%)	

[1] Only includes providers who received all three surveys

[2] Excludes N/A values where demographics were suppressed due to n<=6 for sub-groups.

[3] p-value is Mann-Whitney Wilcoxon or Kruskal-Wallis for IV's relative to total number of responses by clinician

Significance levels: * < .05; ** < .01

Table 4. 4: Response Rates relative to time, day of week and month of survey.

	Total Responses (% Valid)	No Response	% Response Rate	p-value
Month of Survey				**< .001****
Month 1 (August)	303 (89%)	218	58%	
Month 2 (September)	264 (91%)	247	52%	
Month 3 (October)	235 (93%)	287	45%	
Day of the Week Survey Sent				**0.076**
Monday	179 (92%)	131	58%	
Tuesday	158 (93%)	161	50%	
Wednesday	156 (89%)	159	50%	
Thursday	163 (89%)	139	54%	
Friday	147 (90%)	162	48%	
Time of Day				**0.2848**
Early Morning (0600)	256 (93%)	264	49%	
Mid-Day (1100)	281 (89%)	238	54%	
Mid-Afternoon (1500)	266 (91%)	250	52%	

Notes: P-valus associated with chi-square for response or no resonse. With three response categories (Valid, Invalid, and No Response). Month was still highly significant, Day of Week and Time of Day were not

4.3.3 Impact of individual characteristics on burnout response scores

Table 4.5 shows the relationship between individual characteristics and burnout. As hypothesized and as supported by the literature, gender, age, and specialty are all significantly related to burnout. Females were significantly more likely to report feeling some level of burnout, defined as a 3, 4 or 5 response (45% compared to 33% of males). While age group had a significant relationship to burnout scores, this relationship was not linear 40-54 year-olds reporting the highest burnout rates (47%) compared to 32% of 25-39 year-olds and 26% of 55+ year-olds. Specialty group also had a significant relationship to burnout scores with primary care having the highest burnout rate at 42% with surgical specialties close behind at 40%. Non-surgical medical specialties reported lower burnout rates (31%) with emergency medicine reporting the lowest burnout (24%).

Table 4. 5: Burnout scores relative to individual provider specialty and demographics.

Category	Burnout Score 1	2	3	4	5	p-value[2]
Gender[1]						0.002**
Female	44 (14%)	128 (41%)	100 (32%)	24 (8%)	16 (5%)	
Male	76 (22%)	156 (45%)	67 (19%)	31 (9%)	15 (4%)	
Age Group[1]						< .001*
25-39	38 (17%)	112 (50%)	48 (22%)	18 (8%)	6 (3%)	
40-54	47 (15%)	120 (38%)	103 (32%)	26 (8%)	21 (7%)	
55+	56 (34%)	68 (41%)	23 (14%)	14 (8%)	6 (4%)	
Specialty Group						.012*
Primary Care	36 (15%)	101 (42%)	68 (29%)	20 (8%)	13 (5%)	
Non-Surgical Medical	59 (21%)	135 (48%)	50 (18%)	23 (8%)	16 (6%)	
Surgical	43 (26%)	57 (34%)	48 (29%)	13 (8%)	6 (4%)	
Emergency	12 (29%)	19 (46%)	8 (20%)	2 (5%)	0 (%)	

[1] Excludes N/A values where demographics were suppressed due to n<=6 for sub-groups.

[2] p-value is Mann-Whitney Wilcoxon or Kruskal-Wallis for IV's relative to burnout scores

Significance levels: * < .05; ** < .01

4.3.4 Impact of survey administration on burnout scores

Table 4.6 shows the tests of independence of burnout responses from survey characteristics. As hypothesized, none of the survey administration variables impacted the burnout responses. Specifically, neither the time and day of survey administration, nor the day of the week that physicians achieved statistical significance.. A one-way ANOVA test further reinforced the independence of burnout scores from random survey administration day of week (df=4, p-value = .12) and time of day (df = 2, p-value=.95) variables.;). There was no statistically significant relationship of month of survey administration to burnout score (p=.908). Furthermore, physicians' burnout score was not related to the number of times they responded to the survey (1, 2, or 3).

Table 4. 6: Relationship of survey characteristics and burnout responses

Category	1	2	3	4	5	p-value[1]
Month of Survey						**0.908**
Month 1 (August)	56 (21%)	115 (43%)	67 (25%)	21 (8%)	10 (4%)	
Month 2 (September)	50 (21%)	99 (41%)	59 (24%)	21 (9%)	12 (5%)	
Month 3 (October)	44 (20%)	98 (45%)	48 (22%)	16 (7%)	13 (6%)	
Day of the Week Survey Sent						**0.082**
Monday	38 (23%)	71 (43%)	35 (21%)	12 (7%)	8 (5%)	
Tuesday	19 (13%)	70 (48%)	37 (25%)	10 (7%)	11 (7%)	
Wednesday	25 (18%)	68 (48%)	30 (21%)	12 (9%)	6 (4%)	
Thursday	40 (28%)	59 (41%)	29 (20%)	13 (9%)	4 (3%)	
Friday	28 (21%)	44 (33%)	43 (33%)	11 (8%)	6 (5%)	
Time of Day						**0.99**
Early Morning (0600)	52 (22%)	96 (40%)	57 (24%)	22 (9%)	11 (5%)	
Mid-Day (1100)	50 (20%)	108 (43%)	66 (26%)	12 (5%)	14 (6%)	
Mid-Afternoon (1500)	48 (20%)	108 (45%)	51 (21%)	24 (10%)	10 (4%)	
Number of Responses						**0.52**
One	20 (22%)	41 (46%)	21 (23%)	7 (8%)	1 (1%)	
Two	34 (18%)	92 (48%)	47 (24%)	11 (6%)	8 (4%)	
Three	96 (21%)	179 (40%)	106 (24%)	40 (9%)	26 (6%)	
Day of the Week Survey Responded						**0.46**
Monday	29 (21%)	66 (47%)	29 (21%)	9 (6%)	7 (5%)	
Tuesday	26 (20%)	55 (42%)	35 (27%)	7 (5%)	8 (6%)	
Wednesday	30 (19%)	69 (44%)	36 (23%)	13 (8%)	9 (6%)	
Thursday	27 (20%)	57 (43%)	31 (23%)	13 (10%)	5 (4%)	
Friday	26 (20%)	50 (39%)	31 (24%)	15 (12%)	5 (4%)	
Saturday	8 (36%)	9 (41%)	5 (23%)	0 (%)	0 (%)	
Sunday	4 (21%)	6 (32%)	7 (37%)	1 (5%)	1 (5%)	

[1] p-value is Mann-Whitney Wilcoxon or Kruskal-Wallis for IV's relative to burnout scores
Significance levels: * < .05; ** < .01

4.3.4 Test-retest reliability

To test month over month test-retest reliability, we operationalized the data as month, month+1 for any individual respondent who responded in consecutive months and conducted an intra-class correlation. Using the two-way mixed effects model and absolute agreement as recommended by Koo and Li (2016) (Koo & Li, 2016) for both test-retest and intra-rater reliability, the ICC was .871, 95% CI [.843,.894]. This is reinforced by a Wilcoxon signed rank test (p=.307) that we cannot reject the null that the median of the population of differences is zero. Most respondents remained stable month over month (Table 4.7) with all but 4 staying within one of their previous response in consecutive months.

Table 4. 7: Month to month change in provider burnout responses.

Month to Month +1 Change	Total Providers
-2	2
-1	35
0	277
1	45
2	1
4	1

4.3.5: Thematic Analysis of text responses

Respondents were instructed to just respond with the number that corresponds to their level of burnout. However, because of the method of delivery, it was possible for them to provide more information with their response. To account for this additional data, we conducted a thematic analysis of the text response. Overall, half (12) were neutral, 38% (8) were coded as negative, 13% (3) were coded as positive. The positive responses included 2 people who expressed love for their job and one who expressed appreciation for how much the office of "Physician Wellness" had helped them with their burnout. Most of the neutral responses highlighted the challenge of a 5-point scale on a single question to capture burnout. 7 of these respondents wanted to provide more information about the reasons for the score they were submitting, while 5 explicitly rated themselves between two possible responses. Of the negative responses, 2 indicated that we were asking the wrong question, 5 took issue with the delivery of the survey through the EMR, and 3 expressed irritations with the repeatedness of the month over month survey.

4.4 Discussion

The goal of this study was to test the feasibility and reliability of a SIBM administered in the clinical workflow to gather longitudinal measures of physician burnout. The response rate of 58% in the first month is very high when compared to similar surveys. For example, Shanafelt et al (2012) received a 26.7% response rate on a randomly sampled national burnout survey that was e-mailed with three follow-up reminders (Tait D. Shanafelt et al., 2012). Similarly, Olson et al. (2019) e-mailed a link to both the Mini-Z and the MBI to 4118 clinicians in a single AMC. Of those surveyed, 1252 clicked the link (30.4%) and only 557 ultimately responded (13.5%) (Olson, Sinsky, et al., 2019). While these examples are admittedly different with the first being a national sample and both involving more time-consuming surveys, our response rates point to the possibility of this repeated survey method for a pre and post-intervention test. While response rates dropped off by month 3, they remained high at 45% and we would argue that is in part due to the artificial once a month repeated survey associated with this study. There is some indication of response bias with primary care significantly more likely to respond to all surveys than the other specialty groups. While a significance test was not reported, this is potentially different than the Shanafelt et al. in which primary care represented only 26.4% of responders compared to the 38.5% of the general U.S. physician population from which they were randomly sampled. Similarly, there was significant difference in response rate by gender with men more likely than women to not respond at all (34% compared to 24%)

Our findings reinforce the existing literature which indicates that gender, and specialty are all significantly related to burnout. Interestingly, there are some mixed findings on the relationship of age to burnout (Gardner et al., 2019; Kroth et al., 2019; Mehta et al., 2019; E. R. Melnick et al., 2020; Wylie et al., 2014). We also identified a non-linear relationship of age to burnout with burnout highest in the mid-range (40-54) age group. We acknowledge that this may not be consistent with other studies given our exclusion of residents and fellows from the study population. The feasibility of this method is further supported by the fact that the random survey administration variables, which allows us to control for individual level variables, did not have any relationship to burnout scores. The SIBM that we selected has been demonstrated to externally validate to the Emotional Exhaustion dimension of the MBI. Some authors have argued, for example, that Health IT impacts on burnout are characterized more by the depersonalization of medicine which may limit the utility of this SIBM for pre and post testing of Health IT interventions. Other measures such as the SQBS embedded in the mini-Z survey exist that may prove more useful, although a comparison of the two is out of scope of the current study (Olson et al., 2019).

Perhaps most importantly, we find evidence of high reliability of the burnout results from month to month. While we do expect burnout to change over time, both the close time proximity of the test-retest and the fact that there were no apparent macro level interventions, suggest that this method would lend itself to pre- and- post intervention measurement to test the impact of a particular intervention on burnout. The limited individual movement (+1 or -1) on the burnout would reinforces that there is change in burnout sentiment even within a one-month period.

4.5 Limitations

While this current study demonstrates the feasibility and reliability of a SIBM administered in the clinical workflow, there current study has several limitations. We were able to obtain adequate respondent populations to test the current hypotheses, but because this study was conducted at a single AMC, we caution that this limits the generalizability of our findings. Additionally, our decision to exclude residents, fellows and non-physician providers limited our insights into these populations for whom burnout is a serious issue that warrants ongoing study. The current study was conducted between August and October of 2020 during an international COVID-19 pandemic. While we acknowledge the unquestionable impact that the pandemic had on study participants, organizational care delivery and COVID volumes remained steady enough during the study period that the results are still valid, although they may not directly apply to a pre or post-pandemic world.

4.6 Conclusions

Clinician burnout is a known problem in physicians. There are many reasons for the burnout problems beyond the EHR. However, our ability to combat burnout is, in part, dependent on our ability to measure if interventions designed to impact burnout are having the intended effect on the individual level. The current study demonstrates our ability to reliably test and retest physician burnout in a no-cost, low-interruption, privacy protecting fashion. We believe this survey method can be used in the future both to further understand the underlying causes of burnout on a broad scale and to test the impact of future interventions.

Chapter 5: Characterization of EMR usage patterns by specialty and venue

Objective: Electronic Medical Records (EMRs) have been widely identified as a contributor to clerical burden and physician burnout. This study identifies how underlying use patterns vary across specialty and venue to advise future targeted improvement interventions.

Materials and Methods: Event log data from the EMR was captured over a series of three months for a single Midwest Academic Medical Center. Intra-class correlation was used to test the hypothesis that provider EMR usage tends to remain consistent over time and venues. Exploratory factor analysis using ordinary least squares with oblique factor rotation due to assumption of correlation between factors was used to identify latent EMR usage patterns. Regression analysis was then used to identify the extent to which venue and specialty account for the identified latent factors.

Results: All physician event logging metrics had good to excellent ICC (.8 to .96) from month to month, but poor to moderate ICC (.118 to .699) between inpatient and ambulatory. The best fit model overall explained 45% of the variance with six factors across 15 features; for the ambulatory venue explained 52.4% of the variance with 6 factors across 14 features and the best fit model in the inpatient venue explained 71% of the variance with five factors across 13 features. The best single factor was factor 1 for the inpatient model which was comprised of four time in EMR features with factor loadings from .6 to .92 and explained 23% of the total variance. Factors for the overall model were highly significantly related to venue while the ambulatory factors were highly correlated to specialty. This relationship did not hold true for the inpatient model.

Conclusion: Physician EMR usage remains consistent over time for a particular physician practicing in a single venue but varies greatly between venues. Ambulatory EMR usage appears to vary greatly between specialties while this is not the case in the inpatient venue. Future analysis of EMR usage should account for both specialty and venue to account for sub-group differences.

5.1 Introduction

With the now ubiquitous use of electronic medical records in healthcare, attention has turned to the impact that EMR use has on provider well-being. Time spent on "desktop medicine" has grown and is perceived as a barrier to patient communication (Ratanawongsa et al., 2016). Studies using both time-motion and EMR event logging have raised evidence that physicians spend as much time on documentation as they spend with patients (C. Sinsky et al., 2016; Tai-Seale et al., 2017) with one study of EMR interactions finding that physicians spend as much as twice as much time on non face-to-face activities as they spend on patient care (Arndt et al., 2017). The traditional method of time-motion studies of physicians were generally focused on a single site and specialty which limits ability to look at variations across venue and specialty. Increasingly event logging data is being used to understand EMR use, workflow, care team dynamics, and model EMR use and outcomes (Rule et al., 2020).

In one large-scale national descriptive study, Overhage and McCallie (2020) found variation in the average time spent in the EMR by physicians varies across specialties from 7.3 to 22.5 minutes per patient (Overhage & McCallie, 2020). They further found that most time, 74%, was spent on chart review, documentation, and ordering with chart review alone comprising 33% of total time. The study was limited in that it did not attempt to characterize the underlying workflows that comprise of that variation. This current study builds on this evidence by looking at both the consistency of provider EMR over time and between venues using intra-class correlation of EMR usage statistics. We then use exploratory factor analysis (EFA), an unsupervised learning technique to determine the extent to which there are latent EMR practice patterns and identify how they correlate to specialty groups. Our expectation is that individual physician EMR usage will remain relatively consistent over time and between venues, but we expect that overall EMR use patterns will vary greatly between specialties and venue.

5.2 Materials and Methods

This study leverages event log data made available through the EMR vendor at a single Midwest Academic Medical Center. These data consist of logging that activates when the user performs 2 mouse clicks or 15 keystrokes or 1700 mouse miles per minute (Overhage & McCallie, 2020) (Aziz et al., 2019). Data are preprocessed by the vendor and made available to clients through the LightsOnNetwork (https://www.cerner.com/solutions/lights-on-network) and Cerner Advance (https://advance.cerner.com/). Data were obtained for all attending physicians for the months of August, September, and October 2020. The data were categorized and pre-aggregated by provider. For any time variables and limited count variables, data were also collected by venue categorized as ambulatory, inpatient or emergency based on the venue of the patient on whom documentation is occurring. In addition to data broken down by venue, there is an "all" category that is an aggregation of all activity by physician regardless of venue. Prior to the data being made available to the research team, the data were de-identified and individual providers were pre-aggregated into

Table 5. 1: Specialty Groupings

Specialty Group	Specialty
Primary Care	Family Medicine, General Internal Medicine, Pediatrics, Student Health Attending, Women's Health
Surgical Specialties	Acute Care Surgery, General Surgery, Ophthalmology, Orthopaedic, Otolaryngology, Plastic Surgery, Urology
Non-Surgical Medical	Behavioral Health, Cardiology, Dermatology, Endocrinology, Gastroenterology, Internal Medicine (excluding GIM), Nephrology, Neurology, Oncology, Optometrist, PM&R, Pulmonology

specialty groups as shown in Table 5.1. This was necessary as individual specialties were too granular to protect the privacy of the physicians in the study.

A subset of the EMR usage features available were selected based on literature scan, proposed heuristics (C. A. Sinsky et al., 2020) (Baxter, Apathy, Cross, Sinsky, & Hribar, 2021), subject matter expert interview with the Chief Medical Information Officer (CMIO), Associate CMIO, and Medical Director for Decision Support, and data profiling. Initial data was filtered to require > 10 patients seen in a month by the physician. Since time spent in a between venues was continuous, a cutoff a of 25% was used to attribute any physician to a particular venue. Similarly, a cutoff of 35% zero values for any time variables was established for exclusion from analysis. The full variable list, definitions, and normalization applied is available in Appendix B of the supplementary materials. Variables include time in the EMR, including information seeking patterns and documentation outside of core work hours, use of Computerized Physician Order Entry (CPOE), electronic documentation adoption, EMR performance and stability, decision support volume (burden) and interaction, mobility, and use of advanced features designed to assist documentation. For all time and count variables, missing values were imputed as 0 (i.e., no time on that action) and normalized to a per patient value. Count of patients was based on the number of notes signed. Previously derived percentages (CPOE percent, percent of electronic documentation authored, and adoption percent) could not be assumed to be zero, therefore, no values were imputed, and any NULL values were excluded from analyses that included those variables.

To assess month over month ICC for individual providers in a particular venue and across venues, a two-way random-effects Intra-Class Correlation (ICC) model using single rater for consistency was used to test the month over month ICC for individual physicians within a particular venue and between the inpatient and ambulatory venue. (Shrout & Fleiss, 1979) (Koo & Li, 2016). A subsequent exploratory factor analysis was conducted to identify latent patterns of EMR usage. Factor analysis was assessed overall and for the inpatient and ambulatory venues for variables available at that level. The emergency venue was excluded from the factor analysis due to inadequate number of observations. Data were first profiled to identify any problematic partial correlations for underlying variables and feature reduction was considered based on any pairwise correlations > .8 (Murphy, 2021). Kaiser-Meyer-Olkin procedure was performed to determine factor adequacy. Features with the lowest KMO were iteratively removed until the set resulted in an acceptable Tucker-Lewis Index (TLI) of factor reliability of .9. This consistently was a threshold of exclusion of features with a KMO < .75 , a level identified by Kaiser as “middling”. (Cerny & Kaiser, 1977; Dziuban & Shirkey, 1974; Kaiser, 1970). Features where then tested for positive determinants and Bartlett’s test of sphericity to ensure goodness of fit of the factor analysis. The number of features was determined by comparison of both eigenvalues using Kaiser’s Rule and a parallel test.

For the “All” venue, a logistic regression was used to determine the extent to which the latent constructs were representative of variations between ambulatory and inpatient practice patterns. For the Inpatient and Ambulatory venues, logistic regression was used to investigate how latent constructs correlate to each of the specialties. Factor analysis was conducted using an ordinary least squares method with oblique rotation of factors. This rotation method was verified by investigation of correlation between resultant factors which remained high (> .3) across analyses. All analyses were conducted using R version 1.4.1106 libraries IRR for ICC and Psych for factor analysis (Gamer, 2012; Revelle, 2021). Logistic regression was conducted using glm in the base R package. Data and R coding are available from the author on request.

5.2.1 A Note on Outliers

The time in EMR data contained extreme outliers. For example, the variable documentation time per patient is has a standard deviation (SD) of 12.45 . When these data were more closely examined the times were distributed across venue and functions and appeared to be plausibly capturing a physician that is spending an extremely high amount of time on documentation. Given the academic mission of the health system measured, some activities, e.g., extremely high chart review times, may represent extra-clinical activities such as manual chart abstraction. This activity is likely limited by exclusion of residents who likely do most of the manual chart abstraction, but attendings in an AMC may spend time, for example, reviewing resident documentation. The decision was made to retain data for physician EMR use metrics that met the criteria listed in the materials and methods.

5.3 Results

Table 5. 2: Counts of physician-months in each specialty by venue

	Ambulatory	Inpatient	Emergency Department	All venues[1]
Primary Care	342	94		368
Non-Surgical Medical	436	315	2	612
Surgical Specialty	310	140		351
Emergency Specialty	18	63	84	99

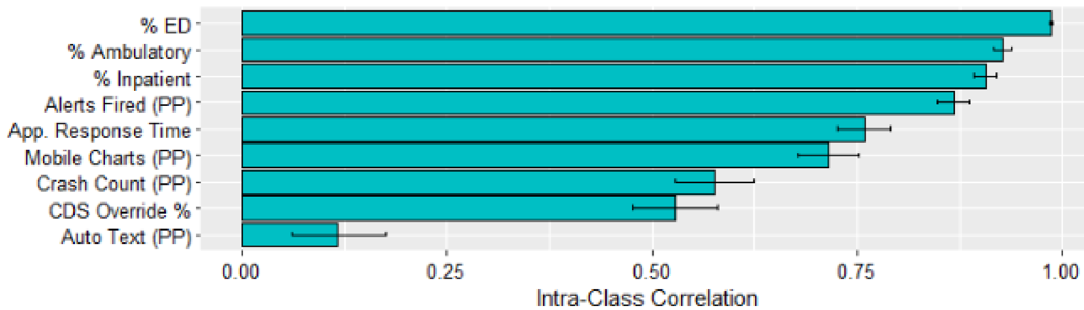
Notes: Includes only physicians with > 25% time in a particular venue
[1] Is an overall total for all venues and includes physicians who have seen > 10 patients for the month

5.3.1 Physician EMR usage month over month and between venues

Table 5.2 shows a count of the physician months that were included in the analysis. Detailed ICC results are available in supplementary materials Appendix E. It was only possible to meaningfully compare the inpatient and ambulatory venues as not enough physicians practice more than 25% in the emergency room and inpatient or ambulatory. Based on the 95% confidence interval, values less than .5, between .5 and .75, between .75 and .9 and greater than .9 should be considered of poor, moderate, good, and excellent reliability, respectively. As expected, physician EMR usage within venue was far more consistent from month to month than was between the ambulatory and inpatient venue across all measures.

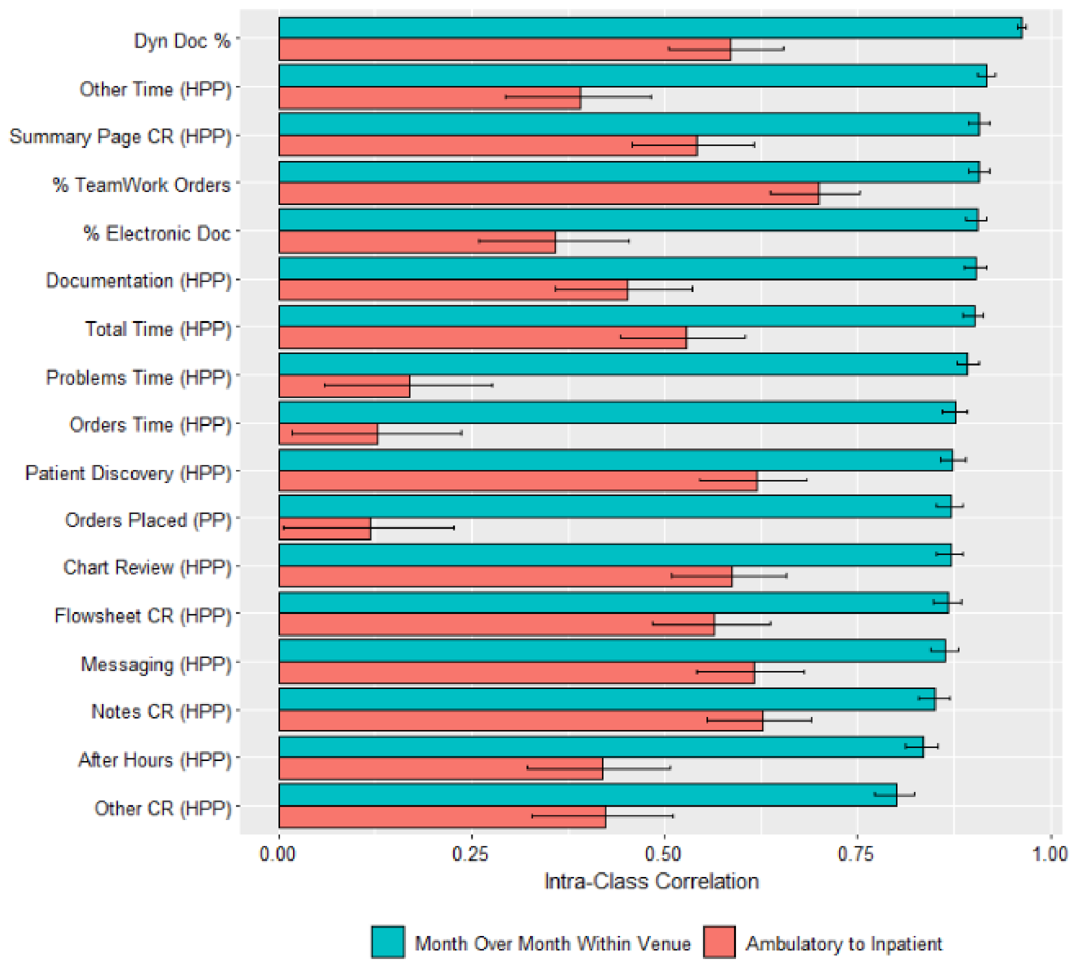
Figure 5.1 shows the overall month over month ICC for each physician. The percent of time spent in ED, Ambulatory or Inpatient remains consistent month over month (ICC=.987; 95% CI[.985,.989]; ICC=.928; 95% CI[.917,.938]; ICC=.906; 95% CI[.892,.92]) respectively. Alerts fired per patients shows good correlation (ICC=.87; 95% CI[.846,.885]), but alert override percentage was low (ICC=.53; 95% CI[.474-.576]). The application experience shows mixed results with application response time experience showing good reliability over time (ICC=.76; 95% CI[.712-.779]), while the number of application crashes is in the moderate range from month to month (ICC=.573;

Figure 5. 1: Month to month ICC for all venues



95% CI[.523-.621]). Of note, Auto Text usage, an advanced templating feature to assist with note creation had very poor reliability (ICC=.114; 95% CI[.058,.173]), suggesting there may be some problem with the underlying data. Figure 5.2 shows Month over month measures consistency of physicians EMR usage within a particular venue from month to month and between ambulatory and inpatient. Almost all usage metrics within venue show as good ICC with two, other time per patient and percent of dynamic documentation showing as excellent month over month (ICC=.918; 95% CI[.869,.949] and ICC=.997; 95% CI[.994,.998] respectively). For those providers who practiced at least 25% of their time in ambulatory and 25% in inpatient, there was poor ICC on most metrics with teamwork for orders remaining the most consistent between venues (ICC=.7; 95% CI[.637,.753]), followed by time spent on reviewing notes, patient discovery and messaging most consistent between venues (ICC=.63; 95% CI[.555-.691]; ICC=.62; 95% CI[.546,.684], and ICC=.62; 95% CI[.529,.571] respectively). All other metrics show poor ICC between venues and several of the usage metrics that showed potential negative ICC may indicate a lack of variation among physicians.

Figure 5. 2: Month over month ICC: Single venue vs inpatient to ambulatory



Notes: Error bars indicate the 95% confidence interval. HPP=Hours Per Patient; PP=Per Patient

5.3.2 Factor analysis of EMR use metrics

All Venues

Table 5.3 shows the results of the exploratory factor analysis for the All Venue after feature reduction was performed using a .3 cutoff for display purposes. Full models are available in the supplementary Appendix C. The 6-factor model achieved an acceptable level of .909 on the TLI, but only accounted for 45% of the variance in the data. Three of the features included did not meet the threshold of .3 loading on any of the factors, other time in EMR, mobile charts opened, CDS overridden, and use of auto-text, indicating that the variance in those features is captured across multiple factors and did not show a strong enough loading to be reported out on any factor. Initial examination of the factors appears to provide evidence that these factors are capturing variability between inpatient and ambulatory workflows. For example, factor 2 is characterized by use of summary pages for information seeking and a high amount of time spent on

Table 5. 3: Factor loadings exceeding a .3 cutoff for all venues

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Total time in EMR						
Chart Review Time	NA[2]	NA[2]	NA[2]	NA[2]	NA[2]	NA[2]
Documentation Time			0.35	0.54		
Prescription Time	0.75					
Message Center Time		0.36				
Patient Discovery Time						0.64
Problems Time		0.59				
Other Time						
Work Outside of Work			0.97			
Computerized Physician Order Entry						
Total Orders	0.87					
Teamwork for Orders	-0.50					
Electronic Documentation Adoption	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]
Decision Support Interaction						
Total Alerts Fired	0.45					0.33
Percent Overriden						
Mobility						
Charts Opened						
Information Seeking Time						
Summary (MPage)		0.91				
Flowsheet					0.53	
Clinical Notes				0.75		
Other					0.62	
Advanced Features						
Dynamic Documentation %	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]
Auto-text						
Variance Accounted For	**0.115**	**0.092**	**0.074**	**0.064**	**0.06**	**0.047**

[1] All counts and times are normalized as per patient. TLI for model = .909 and Cumulative Variance=.45

[2] Excluded due to high correlation to Information Seeking features

[3] Excluded due to low or middling KMO score (Dyn. Doc.=.59 and Electronic Doc. Adoption=.71)

Table 5. 4: Regression of all venue factors on flag for ambulatory and inpatient physicians

	Ambulatory			Inpatient			Odds of Practice in Venue	
	Estimate	Odds Ratio	P-Value	Estimate	Odds Ratio	P-Value	Ambulatory	Inpatient
Factor 1	0.33*	1.49	0.02	-0.03	0.98	0.84	Increase Odds	NS
Factor 2	1.49***	4.43	< 2E-16	-1.29***	0.28	5.20E-15	Increase Odds	Decrease Odds
Factor 3	-1.78***	0.17	< 2E-16	0.75***	2.12	1.14E-11	Decrease Odds	Increase Odds
Factor 4	3.03***	20.68	< 2E-16	-1.01***	0.37	3.65E-09	Increase Odds	Decrease Odds
Factor 5	-2.28***	0.1	< 2E-16	1.44***	4.22	2.49E-15	Decrease Odds	Increase odds
Factor 6	-0.65***	0.52	3.16E-06	0.14	1.15	0.224	Decrease Odds	NS

Significance Levels: *=.05; **=.01;***=.001
NS = Not Significant

managing the problem list and message center. Activities that are generally seen in ambulatory settings. Conversely, factor 3 is primarily characterized by high work outside of work. This is expected to be almost entirely driven by inpatient since our after-hours measure (outside of 6:00am to 6:00 pm) fails to capture shift work accurately. Table 5.4 shows how each of these factors performs in a logistic regression using binary inpatient and ambulatory designations defined as greater than 50% of patients seen in inpatient and ambulatory venues, respectively and table six summarizes the impact on odds of ambulatory and inpatient for significant factors. Factors 2-5 are highly correlated to both the inpatient and ambulatory venues, with evidence that factor 1 and factor 6 are correlated with ambulatory, but not with inpatient. Factors 1, 2, and 4 increase the odds that a provider practices in the ambulatory venue while factors 3,5 and 6 are decrease the odds of ambulatory practice. Conversely, factors 2 and 4 decrease the odds that a provider is inpatient while factors 3 and 5 increase the odds. Similarly, Table 5.5 shows how the factors relate to specialty and a summary of the directionality of significant relationships between factors and specialty is in Table 5.6.

Table 5. 5: Regression of overall EMR use factors on specialty

	Primary Care			Non-Surgical Medical			Surgical			Emergency		
	Estimate	Odds Ratio	P-Value	Estimate	Odds Ratio	P-Value	Estimate	Odds Ratio	P-Value	Estimate	Odds Ratio	P-Value
Factor 1	.24*	1.27	0.022	.20*	1.22	0.019	-.22*	0.8	0.043	-2.04***	1.3	1.46E-10
Factor 2	1.46***	4.32	< 2E-16	-.98***	0.38	< 2E-16	-.51***	0.6	0.000	.68*	1.97	2.28E-02
Factor 3	-.47***	0.63	8.25E-05	-.13	0.88	0.134	-.32*	0.73	0.014	2.38***	1.08	2.71E-13
Factor 4	-.75***	0.47	3.75E-06	1.18***	3.24	< 2E-16	-.20	0.82	0.150	-7.57***	5.16	< 2E-16
Factor 5	.66***	1.94	1.66E-05	.38**	1.46	0.008	-2.15***	0.12	< 2E-16	.01	1.01	0.9782
Factor 6	-1.40***	0.25	< 2E-16	-0.15	0.86	0.087	1.38***	3.96	< 2E-16	1.2***	3.32	3.10E-05

Significance Levels: *=.05; **=.01;***=.001

NS = Not Significant

Table 5. 6: Significant relationships of all venue factors with specialty groups

	Odds of Specialty			
	Primary Care	Non-Surgical Medical	Surgical	Emergency
Factor 1	Increase	Increase	Decrease	Increase
Factor 2	Increase	Decrease	Decrease	Increase
Factor 3	Decrease	NS	Decrease	Increase
Factor 4	Decrease	Increase	NS	Increase
Factor 5	Increase	Increase	Decrease	NS
Factor 6	Decrease	NS	Increase	Increase

NS = Not Significant

Ambulatory Venues

As with the overall model, Actual time, Chart Review Time, and Adoption Percentage were excluded due to high correlation with other variables in the model. Dynamic Documentation percentage was excluded due to low individual feature KMO (.56). As mentioned above, CDS, mobility, application performance and advanced features (auto text) which do not vary across venues were excluded from this analysis. The overall KMO of the features included was .84. Eigenvalue inspection recommends a 6-factor model while parallel analysis recommends 5. The 6-factor had an acceptable TLI of .909 and explains 52% of the total variance. The resultant factors are in Table 5.7 for all loadings > .3. Subsequent regression on the four specialty groupings in in Table 5.8 with a summary of directionality of significant relationships in Table 5.9.

Table 5. 7: Factor loadings exceeding a .3 cutoff for ambulatory venue

Total time in EMR	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Chart Review Time	NA[2]	NA[2]	NA[2]	NA[2]	NA[2]	
Documentation Time			0.65			
Prescription Time		0.72				
Message Center Time		0.43				
Patient Discovery Time	0.96					
Problems Time		0.52		0.40		
Other Time	0.36			0.36		
Work Outside of Work			0.77			
Computerized Physician Order Entry						
Total Orders		0.49			0.37	
Teamwork for Orders					-0.83	
Electronic Documentation Adoption						0.67
Information Seeking Time						
Summary (MPage)				0.78		
Flowsheet	0.41	0.36				
Clinical Notes	0.63		0.30			
Other					0.31	
Advanced Features						
Dynamic Documentation %	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]
Variance Accounted For:	**0.128**	**0.1**	**0.09**	**0.08**	**0.073**	**0.05**

[1] All counts and times are normalized as per patient. TLI for model = .904 and Cumulative Variance=.47

[2] Excluded due to high correlation to Information Seeking features

[3] Excluded based on KMO score (Dyn. Doc.=.56)

Table 5. 8: Regression of ambulatory EMR use factors on specialty

	Primary Care			Non-Surgical Medical			Surgical			Emergency		
	Estimate	Odds Ratio	P-Value	Estimate	Odds Ratio	P-Value	Estimate	Odds Ratio	P-Value	Estimate	Odds Ratio	P-Value
Factor 1	-1.39***	0.25	< 2E-16	.52***	1.68	2.36E-06	.44**	1.55	0.003	.74*	2.1	0.049
Factor 2	.997***	2.71	2.19E-10	-.91***	0.4	5.19E-11	-1.12***	0.32	3.02E-05	1.42***	4.12	0.0008
Factor 3	-.50***	0.61	0.001	.56***	1.75	1.47E-05	-1.08***	0.34	2.53E-05	-2.33***	0.1	0.0009
Factor 4	1.55***	4.71	< 2E-16	-10.4***	0.35	< 2E-16	-.55**	0.57	0.0060	-.64	0.53	0.086
Factor 5	.97***	2.63	2.96E-08	1.38***	3.97	< 2E-16	-1.88***	0.15	< 2E-16	-1.81***	0.16	4.60E-05
Factor 6	-.50***	0.61	0.0001	-.22*	0.8	0.043	.82***	2.26	2.17E-06	2.86***	17.44	4.06E-06

Significance Levels: *=.05; **=.01;***=.001

NS = Not Significant

Table 5. 9: Significant relationships of ambulatory venue factors with specialty groups

	Odds of Specialty			
	Primary Care	Non-Surgical Medical	Surgical	Emergency
Factor 1	Decrease	Increase	Increase	Increase
Factor 2	Increase	Decrease	Decrease	Increase
Factor 3	Decrease	Increase	Decrease	Decrease
Factor 4	Increase	Decrease	Decrease	NS
Factor 5	Increase	Increase	Decrease	Decrease
Factor 6	Decrease	Decrease	Increase	Increase

NS = Not Significant

Inpatient Venue

Once again, Actual time, Chart Review Time and Adoption Percentage were excluded due to high correlation with other variables in the model. Dynamic Documentation percentage was again excluded due to low individual feature KMO (.65) as was messaging time (.57) leaving an overall KMO MSA of .83. As mentioned above, CDS, mobility, application performance and advanced features (auto text) which do not vary across venues were excluded from this analysis. Kaiser's rule recommends a minimum of 6 factors, while parallel analysis recommends 4. 5 factors gives a TLI of .906 and the 6th factor only explains an additional 4% of the variance (Table 5.10). The 5-factor model is below. Factor regression on specialty is in Table 5.11 and directionality is in Table 5.12. Notably, almost none of the inpatient factors had a significant relationship to specialty.

Table 5. 10: Factor loadings exceeding a .3 cutoff for inpatient venue

Total time in EMR	**Factor 1**	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**
Chart Review Time	NA[2]	NA[2]	NA[2]	NA[2]	NA[2]
Documentation Time			0.94		
Prescription Time		0.70			
Message Center Time	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]
Patient Discovery Time	0.60				
Problems Time		0.35		0.45	
Other Time				0.37	0.32
Work Outside of Work			0.42		0.57
Computerized Physician Order Entry					
Total Orders		0.88			
Teamwork for Orders		-0.47			
Electronic Documentation Adoption		0.66			
Information Seeking Time					
Summary (MPage)				0.72	
Flowsheet	0.79				
Clinical Notes	0.92				
Other	0.77				
Advanced Features					
Dynamic Documentation %	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]
Variance Accounted For:	**0.23**	**0.19**	**0.11**	**0.11**	**0.07**

[1] All counts and times are normalized as per patient. TLI for model = .904 and Cumulative Variance=.71

[2] Excluded due to high correlation to Information Seeking features

[3] Excluded based on KMO score (Messaging Time=.57,Dyn Doc=.65)

Table 5. 11: Regression inpatient EMR use factors on specialty

	Primary Care			Non-Surgical Medical			Surgical			Emergency		
	Estimate	Odds Ratio	P-Value	Estimate	Odds Ratio	P-Value	Estimate	Odds Ratio	P-Value	Estimate	Odds Ratio	P-Value
Factor 1	-.25	0.78	0.358	.74***	2.1	4.13E-05	-.32	0.73	0.110	-2.36***	0.09	4.00E-04
Factor 2	-.36	0.7	0.154	-.23	0.8	0.075	.87***	2.38	1.62E-08	-5.70***	0.003	3.94E-06
Factor 3	-.57*	0.57	0.032	.26	1.29	0.061	.14	1.15	0.309	-2.89**	0.06	1.00E-03
Factor 4	.20	1.21	0.398	-.09	0.91	0.510	-.24	0.79	0.168	1.23	3.41	0.057
Factor 5	-.16	0.85	0.491	-.16	0.85	0.171	-.38**	0.68	0.005	3.07***	21.5	3.96E-09

Significance Levels: *=.05; **=.01;***=.001

Table 5. 12: Significant relationships of inpatient venue factors with specialty groups

	Odds of Specialty			
	Primary Care	Non-Surgical Medical	Surgical	Emergency
Factor 1	NS	Increase	NS	Decrease
Factor 2	NS	NS	Increase	Decrease
Factor 3	Decrease	NS	NS	Decrease
Factor 4	NS	NS	NS	NS
Factor 5	NS	NS	Decrease	Increase

NS = Not Significant

5.4 Discussion:

The primary purpose of this analysis was to identify practice patterns that vary between venue and specialty to advise future sub-group analysis. The ICC analysis reinforces that while a particular physician's EMR usage shows good or excellent month over month reliability *within a particular venue*, this does not hold true when practicing across the inpatient and ambulatory venues with ICC's ranging only from poor to moderate. Based on observation and optimization that has been done across different venues in the EMR, we expect that information seeking would vary across venues with outpatient tending to use the summary pages while inpatient appears to spend more time reviewing notes and flowsheets. Both information seeking patterns show moderate reliability between the outpatient and inpatient venue, suggesting that physicians tend to navigate chart review in the same way regardless of venue. Total orders and the time spent on orders show virtually no ICC between venues (.120 and .188 respectively). It's important to note that important confounders, such as patient acuity, are not captured here making it impossible to determine what underlying causes are driving this difference in practice. It is also noteworthy that the 95% confidence intervals for most of the month over month ICCs were small compared to the venue-to-venue comparison, providing further evidence of consistent EMR usage patterns month over month. ICC also revealed that the provider's time in the various venues (inpatient, ambulatory and emergency) remains stable over time with all ICCs in the > .9 range. Application performance and stability showed good reliability month over month, suggesting that there is some variance in the user experience. The total alerts fired approached excellent month over month ICC (.868), but surprisingly the override percentage barely met the good reliability threshold (.528). This suggests that while providers may encounter the same alert burden month over month, they are not as consistent in override behavior.

Our expectation that factor analysis across all venues would identify latent patterns that distinguish inpatient from outpatient practice was confirmed; each factor that significantly correlates to ambulatory and inpatient practice in a logistic regression has opposite odds for each venue (e.g., increased odds for ambulatory and decreased odds for inpatient or vice versa). Factor 1 from the "all venue" analysis is EMR usage characterized by high total orders and orders time, but low teamwork for orders. Somewhat counterintuitively, given the acuity of inpatients relative to outpatients, this factor increases the odds that a physician is ambulatory and was not significantly related to inpatient practice. Factor 2 in the "all venue" high usage of summary pages for chart review as well as high messaging and management of the problem list. These are all activities that are generally associated with ambulatory primary care physicians, which is confirmed by the logistic regression results. Factor 3 is driven by high work outside of the hours of 6:00 and 6:00 which is just indicative of “shift work” as demonstrated by the fact that this increases the likelihood that a physician is either inpatient or an emergency specialty, both of which are characterized by high shift work. Factor 4 is driven by time in clinical notes both for documentation and chart review. This is associated with ambulatory physicians and specifically increases the likelihood that a physician is non-surgical medical and decreases the odds a physician is primary care. Conversely, factor 5 is characterized by chart review time in the flowsheet and “other” which is behavior indicative of inpatient. It’s important to note that none of the factors explains a high amount of total variance with only 45% of the variance accounted for by all six factors. The understanding of the latent practice patterns becomes a little more informative in-venue. For those physicians practicing in the ambulatory venue, time in the flowsheet and clinical notes (factor 1) is increases the odds of all specialties except for primary care. Conversely, time in the summary pages along with problems time (factor 4) increases the odds of primary care. Factor 2, which increases the odds of primary care and ED, is characterized by high problems, messaging, and prescription time as well as high teamwork for orders, practice which is frequently associated with primary care. Somewhat surprisingly, factor 3, high documentation time and work outside of work, increases the odds that a physician is in a non-surgical medical specialty and decreases the odds that they are in primary care or a surgical specialty.

In contrast to the ambulatory venue in which the factors highly correlated with the specific specialties, the inpatient factors did not appear to be a proxy for specialty. This suggests that EMR usage is more homogenous across specialties in the inpatient setting than in the outpatient setting. 5 factors also accounted for 71% of the variance suggesting that these more truly represent latent differences in EMR usage in the inpatient setting. In the inpatient venue, almost none of the factors were associated with primary care, except for high documentation and shift work (factor 3), which decreased the likelihood that a physician is in primary care. High chart review time in flowsheet, clinical notes and “other” (factor 1) is associated with non-surgical medical specialties. High prescription time and total orders (factor 2) increases the likelihood of a surgical specialty but was not significant in distinguishing between primary care and non-surgical medical specialties. Given the higher number of these specialties, this finding likely reinforces that their EMR usage is similar in the inpatient setting. Use of summary pages and the problem list (factor 4) was not significant for any of the specialties. Given the low variation captured by the individual factors, we stop short of putting any of these forward as meaningful latent workflows. Given the data currently available, a recent nominal logistic regression approach to understanding how time in EMR relates to specialty may provide more interpretable insight (Wilkinson et al., 2021).

Overall, both the ICC and the factor analysis show varying EMR usage patterns between the venue, suggesting that any analysis of EMR usage should consider sub-group variation within venues and specialties. Furthermore, the factor analysis shows high heterogeneity of EMR usage between the specialties in the ambulatory setting, a finding that is less apparent in the inpatient setting suggesting that while analysis of inpatient EMR usage may be appropriate, it is likely necessary to break the ambulatory down specialty by specialty to uncover sub-group signals that attenuate at the aggregate.

This analysis depends entirely on existing EMR use metrics as provided and validated by the EMR vendor, which likely limits utility. Improvements are underway including heuristics of commonly defined measures to characterize EMR use (C. A. Sinsky et al., 2020) and efforts to standardize measures across EMR vendors (Edward R. Melnick et al., 2021). We believe this needs to continue to be an active area of research including need ongoing direct observation studies to understand and characterize the different kinds of EMR usage and strategies to improve the EMR experience of extreme users. While Sinsky et al.'s (2020) framework for understanding EMR usage is helpful, the metrics remain high-level. Future work to use event log data to characterize specific problematic actions (e.g., constant refresh of lab results), capture networks of communication and understand teamwork, and define different pathways through the application will begin to improve our ability to differentiate patters of EMR usage. Given the investment in Health IT improvement, future work should include the ability to capture advanced informatics and workflow improvements such as use of scribes, transcription, and mobility.

5.5 Limitations

This study examines only a single Academic Medical Center and therefore is not generalizable to the broader physician population. Further, this study leverages pre-existing EMR usage metrics provided by the vendor. While these measures have, in part, been validated elsewhere through both time-motion and observational studies, to do so is out of scope of this analysis, which poses several challenges. The data are monthly aggregates and therefore don't capture more granular variations over time. The primary "per patient" denominator was operationalized by the EMR vendor as the number of notes signed in a particular day. Since physicians do not always sign notes on the same day that a patient is seen, this has limitations, but the author believes that this is likely less of a problem in the monthly aggregate. Certain variables that are theoretically meaningful ended up being excluded: for example, health maintenance time and use of mobile platforms are expected to differentiate between both venue and specialty but had to largely be excluded due to missingness. Other variables were not captured at the individual venue and thus had to be excluded from those analyses. Some variables, such as use of the shortcut functionality, auto-text displayed unexplainable variations across time. While the latent constructs uncovered in the factor analysis were unclear and likely not actionable beyond the scope of this analysis, the author believes that these limitations are acceptable since the intent of this study was to look at the implications of between specialty and between venue variations in EMR usage. Finally, while the author believes that factor analysis at the specific specialty and venue level would reveal more clear latent constructs, the number of observations (physicians) did not allow for that level of analysis.

5.6 Conclusion

This study represents a novel approach to leveraging event logging data to understand underlying EMR usage patterns across venue and specialty. Physician EMR usage remains consistent over time for a particular physician practicing in a single venue but varies greatly between venues. Furthermore, ambulatory EMR usage appears to vary greatly between specialties while this is less the case in the inpatient venue. Future analysis of EMR usage should account for both specialty and venue to account for sub-group differences.

Chapter 6: Single population correlation of EMR metrics to SIBM

Background: Electronic Medical Record (EMR) use, and the associated documentation burden has been shown to contribute to burnout in physicians. The rise of ubiquitous use of EMRs over the last decade along with regulatorily-mandated event logging presents an opportunity to leverage these electronic tools to scale up an EMR-based system for measuring and monitoring the impact of improvement efforts in a learning health system.

Objective: Demonstrate the ability to correlate EMR use to burnout using an EMR-based measurement system comprised of an externally validated Single-Item Burnout Measure (SIBM) and event logging to characterize physician EMR use.

Materials and Methods: From August to October of 2020 an SIBM was administered once a month at randomized times and days of the week to all faculty physicians at an academic medical center. Monthly vendor-provided EMR activity logging was collected during this same period for the same cohort of physicians. ICC was used to test the level of consistency for physician measures on both the independent (IV) and dependent (DV) variables. DV and IVs were averaged across the months and univariate logistic regression was used to identify underlying correlation to a binary self-reported burnout measure across all physicians overall and for individual venues (ambulatory, inpatient and emergency). Multivariate logistic regression was used to test the strength of the relationship when controlling for individual gender, age, and specialty.

Results: Burnout and EMR usage measures showed good to excellent month over month ICC (ICC > .75) except for use of mobile platforms (ICC=.715; 95% CI[.677,.751]), application stability (crashes) (ICC=..573; 95% CI[.523-.621]) which were in the

moderate range (.5 < ICC < .75), decision support overrides which failed to reach the moderate range at the 95% confidence level (ICC=.53; 95% CI[.474-.576]) and use of the advanced autotext feature which had poor ICC (ICC=.114; 95% CI[.058,.173]). The overall model shows high correlation between gender, age-range, and the non-surgical medical specialty relative to primary care across all EMR use measures. In the overall model burnout correlates strongly with percent of ambulatory time (OR=2.4, p=.008) and inversely with percent of inpatient time (OR=.48, p=.053). Total time (OR=1.02, p=05), orders time (OR=1.31, p=.003), total orders (OR=1.12, p=.032), and problems time (OR=5.78, p=.007) were all significantly related to burnout with chart review only significant after controlling for individual level variables (OR=1.1, p=.026). Work outside of work was significant across ambulatory, inpatient, and all venues (OR=1.08, p=.054; OR=1.09, p=.05; OR1.07, p=.02). The results for the ambulatory model are like the overall model except for total time in EMR which was not significantly related to burnout (OR1.01, p=3.47) in the ambulatory venue. EMR variables had almost no correlation to burnout at the .05 level in the inpatient and emergency venues.

Conclusion: This study demonstrates the utility of an EMR-based model for repeated measures of EMR contributions to burnout and scale-up of the measurement of impact of interventions within a Learning Health System (LHS). Our results reinforce existing evidence on the relationship of individual-level characteristics and EMR usage with burnout. To apply this system to pre-post intervention measurement of interventions to reduce EMR causes of burnout, future work should focus on enhancing how we use event log data to measure relevant EMR activity and scale up, possibly across sites, in a randomized fashion to demonstrate causal impact of interventions to reduce EMR-related physician burnout.

6.1 Introduction and Background:

The 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act drove electronic medical record adoption to 96% of acute care hospitals and 86% of physician offices by 2017 through various programs that initially provided incentives for EMR use and later levied penalties for non-adoption (J. Adler-Milstein & Jha, 2017)(Office of the National Coordinator for Health Information Technology, 2019; 2020) . This led to a subsequent rise in the recognition of how documentation burden, driven both by legislation and EMRs, contributes to the phenomenon of physician burnout (Yan et al., 2021; Nguyen et al., 2021). Increasing awareness of the impact of physician wellness has led to calls to address clinician well-being as the fourth pillar of the Quadruple Aim (O'Connor, 2015) and led to calls to reduce documentation burden by 75% by 2025 (Rossetti et al.). Achieving this goal will require action at all levels including Health Information Technology vendors, providers and health systems, and policy. Dymek et al. provide evidence for multiple-interventions address 3 of the highest contributors to EMR burnout: documentation, chart review, and inbox management (Dymek et al., 2021). While Chapter 2 identified an increase in evidence about the link, between EMR use and burnout, most current work is observational; only 3 of 27 (11%) of studies included an intervention with pre-post measurement of intervention effect. By contrast, in an informatics-informed learning health system, improvements are regularly made to improve the physician EMR experience including just-in-time training, EMR advanced feature innovation, and configuration. Scale up of successful interventions to reduce EMR contributions to burnout necessitates that we have a way to reliably collect repeated measures to quantify the impact of improvement efforts.

In the last 15 years, our ability to capture EMR usage activity through event logging has become ubiquitous as vendors and organizations have had to work to become compliant with the 2005 Security Rule of the Health Insurance Portability and Accountability Act (HIPAA). This rule mandated that healthcare organizations record and examine any access activity to Protected Health Information (PHI). This was followed by the 2014 Stage II Meaningful Use regulations which clarified the standards to which those logs must adhere. (Hash, Bowen, Johnson, Smith, & Steinberg, 2005). Because of this, almost all EMRs now track who accessed the patient record, at what time and what action they performed on that record. Over time, event logging data has been increasingly used for many purposes including: counts of actions; counts of higher-level activities; activity durations, clusters, and sequences; and development EMR user networks (Rule et al., 2020), with some researchers suggesting that these data can be more widely used for Health Services Research (Adler-Milstein, Adelman, Tai-Seale, Patel, & Dymek, 2020). In this chapter we examine the relationship between EMR use and burnout as measured by an EMR-based system and consider the implications for future EMR-based learning health systems.

6.2 Materials and Methods:

From August to October 2020 a single item burnout measure was deployed in the EMR message center at a single Midwest Academic Medical Center. The question asks "Overall, based on your definition of burnout, how would you rate your level of burnout? 1) I enjoy my work, I have no symptoms of burnout; 2) Occasionally I am under stress, and I don’t always have as much energy as I once did, but I don’t feel burned out; 3) I am definitely burning out and I have one or more symptoms of burnout, such as physical and emotional exhaustion; 4) The symptoms of burnout that I’m experiencing won’t go away. I think about frustration at work a lot; 5) I feel completely burned out and often wonder if I can go on. I am at the point where I may need to make changes or may need some help.” This question has been previous shown to externally validate to the Maslach Burnout Inventory Emotional Exhaustion (MBI:EE) sub-scale (Rohland et al., 2004), and details of the survey development, administration, feasibility and reliability, and privacy considerations are available elsewhere (Wilkinson, Chapter 4). During this same period monthly event logging data were provided by the EMR vendor that capture: time in EMR, both as a raw number and a percent of total for each activity; time on different methods of information seeking (chart review); work outside of work; orders burden; decision support burden; mobility; and use of advanced features. Details on collection and development of these measures and validation of time metrics are available elsewhere (Overhage & McCallie, 2020; Aziz et al., 2019). Data are available pre-aggregated by month and provider and made available to clients through the LightsOnNetwork (https://www.cerner.com/solutions/lights-on-network) and Cerner Advance (https://advance.cerner.com/). A subset of the available EMR usage features were selected based on literature scan, proposed heuristics (C. A. Sinsky et al., 2020; Baxter et al., 2021), subject matter expert interview with the Chief Medical Information Officer (CMIO), Associate CMIO, and Medical Director for Decision Support, and data profiling. Any physicians with < 10 total patients in a month were excluded and any variables with > 35% missingness or zeros were excluded. For any time in EMR variables, data were further broken down by venue (inpatient, ambulatory, emergency, and an aggregate “all” venue). Data were compiled and provided by an honest broker. Individual level characteristics: gender, age range, and specialty group, were provided for all groupings that contained more than 6 physicians and suppressed for groupings of fewer than 6 in accordance with the approved de-identification protocol. Details on definitions of specialty groupings is available in Appendix D and definitions of event logging variables is available in Appendix B of supplementary materials.

6.2.1 Data Preparation

Raw data were provided including total time of identified activities, counts of events (e.g., decision support alerts, orders place, charts opened on a mobile platform), or pre-derived percentages (e.g., level of teamwork for orders and electronic documentation adoption). Counts of the total patients seen were also provided based on the number of notes signed by that physician in a particular month. For all time and count variables, missingness was interpreted as “none” and 0 was imputed. These variables were converted to total minutes and were subsequently normalized per patient. Frustration with time in EMR can be conceptualized based on the total actual time as well as the percent of time on a particular activity vs another activity. For example, two physicians may spend a high amount of time per patient on Computerized Physician Order Entry (CPOE), but that may account for a small percentage of one physician’s total time and a high percentage of another physician’s time, signaling that orders accounts for a disproportionate burden on the second physician. Because of this, the major categories of time (chart review, documentation, orders (prescription time) message center, problems management and other, were considered both as time per patient and as percent of total time. For any pre-derived percentages, missingness could not be assumed to be zero and therefore any missing values were dropped from those relevant analyses. A minimum cutoff of 25% of activity in a particular venue was used to attribute physicians to ambulatory or inpatient. All Emergency physicians meet this 25% threshold and were attributed to the Emergency Department venue.

6.2.2 Regression analysis

The outcome variable is skewed to the right due to a mode response of 2 (312 out of 729 or 43%) and 63% of responses (462/729) of either 1 or 2, the dependent variable is operationalized as burnout with 1 or 2 indicating "no burnout" and 3,4, and 5 indicating that the respondent is burning out or burned out. This approach is commonly taken among other analysis on SIBMs. Because the data represent repeated measures, a standard logistic regression is problematic due to the introduction of repeated clinician measures as a random effect that needs to be accounted for. Because of this, both the dependent variable (DV), self-reported burnout and independent variables (IV) were operationalized as an average value for each clinician. To validate this approach, we examined intra-class correlation (ICC) from month to month for physician burnout and across all three months within venue for IVs. Because of the high collinearity between the EMR usage variables, a univariate regression for each IV was conducted. Subsequent analyses were run controlling for individual demographic attributes of gender, specialty group and age. Regression analyses were calculated for the all venue, inpatient and ambulatory venues as defined the > 25% cutoff, and ED was restricted to just qualifying Emergency Physicians. All ICCs were calculated as two-way random effects models with single rater for consistency using R Studio version 1.4.1106 libraries IRR for ICC

(Gamer, et al. 2019). Logistic regression was conducted using glm in the base R package. Data and R coding can be made available from the author on request.

6.3 Results

Full ICC results are reported in supplementary materials Appendix E. As a rule of thumb, ICCs greater than .9 indicate excellent reliability, between .75 and .9 indicates good reliability, .5 to .75 indicates moderate reliability, and anything under .5 indicates poor reliability (Koo & Li, 2016). Generally, the approach of averaging across months to account for the random effect is supported by good to excellent month to month ICC values for the burnout score (ICC=.871; 95% CI[.843,.894]). ICCs were good to excellent for IVs with the exception of use of mobile platforms (ICC=.715; 95% CI[.677,.751]), application stability (crashes) (ICC=..573; 95% CI[.523-.621]) which were in the moderate range (.5 < ICC < .75), decision support overrides which failed to reach the moderate range at the 95% confidence level (ICC=.53; 95% CI[.474-.576]) and use of the advanced autotext feature which had poor ICC (ICC=.114; 95% CI[.058,.173]).

Univariate regression models are reported in Table 6.1. While burnout overall was correlated to more EMR use variables than within any individual venue, the percent of time in each venue was also correlated to burnout, with ambulatory highly positively correlated to burnout (OR=2.4, p=.008). Percent of inpatient had an Odds Ratio (OR) < 1, indicating that physicians with higher percent of time in inpatient are less likely to report burnout, although at a lower significance threshold (OR=.48, p=.053). Almost all time in EMR variables were associated with burnout, except for documentation and "other" time. Orders time per patient and problems time per patient were most highly positively correlated to burnout (OR=1.31, p=.003 and OR=5.78, p=.007 respectively). Interestingly, those two as a percent of total time were the only two that show a significant correlation to burnout (OR=521.51, p=.003 and OR=1.57E+09, p=.052), suggesting, for example, that not only was overall time on orders related to burnout, but that a disproportionate amount of time on orders as a percent of total was also related to burnout, although this could not be confirmed through multivariate analysis. The extremely high odds ratios indicated that one unit increase (in this case, 100%) would have a very high impact on burnout, although in practice none of these activities could ever be increased by a unit of 100%. Total orders per patient was also correlated to burnout (OR=1.12, p=.032). Work outside of work, defined as EMR documentation outside the hours of 6:00am and 6:00 pm, was significant in all but the Emergency model (OR1.07, p=.02; OR=1.08, p=.054; OR=1.09, p=.05 for all, ambulatory and inpatient respectively). The model for ambulatory is most like the overall model with the noteworthy exception that total time in EMR did not approach significance for our cohort. EMR variables had almost no correlation to burnout at the .05 level in the inpatient and emergency venues except for work outside of work for inpatient, as reported above.

Table 6.2 shows the same results of the individual regressions of EMR usage variables on burnout while controlling for gender, age, and specialty. For categorical variables, the omitted variables (female for gender, primary care for specialty group, and < 40 for age) served as reference categories for the regression. Male gender correlated negatively with burnout, as did non-surgical medical specialty. Mid-career (age 40-45) had a positive relationship to burnout as compared to the control of < 40. The relationship of EMR usage to burnout remains relatively consistent even when controlling for demographics. The noteworthy exceptions are that message center time per patient overall, while still achieving 90% significance, fell well below the .05 level while messaging as a percent of total time was significant at the 95% confidence level (OR=.0005, p=.05). Orders remained highly significant across total orders per patient, total time on orders, and orders as a percent of total time (OR=1.14, p=.019;

Table 6. 1: Univariate regression results overall and by Venue

	All (n=316)		Ambulatory (n=258)		Inpatient (n=144)		Emergency[1] (n=19)	
Measure	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**
Total time in EMR Per Patient	1.02**	0.040	1.01	0.347	1.03	0.191	1.27	0.191
Chart Review Time Per Patient	1.06*	0.074	1.04	0.295	1.04	0.375	1.32	0.396
Documentation Time Per Patient	1.02	0.389	0.98	0.512	1.09	0.109	1.28	0.649
Prescription Time Per Patient	1.31***	0.003	1.23**	0.023	1.16	0.344	0.45	0.670
Message Center Time Per Patient	1.54**	0.014	1.53**	0.024	1.43	0.375	27.45	0.191
Problems Time Per Patient	5.78***	0.007	3.51**	0.037	0.96	0.988	0.00	0.498
Other Time Per Patient	1.24	0.395	2.47*	0.088	1.23	0.537	1.90	0.148
Information Seeking								
Summary (MPage) CR Time Per Patient	1.05	0.569	1.03	0.717	0.91	0.642	18.18	0.105
Flowsheet Cer Time Per Patient	1.10	0.352	1.02	0.879	1.18	0.175	1.17	0.790
Clinical Notes CR Time Per Patient	1.16*	0.053	1.09	0.281	1.12	0.422	1.89	0.514
Other CR Time Per Patient	1.62*	0.050	1.90**	0.044	1.23	0.371	904.99	0.385
Work Outside of Work Per Patient	1.07**	0.020	1.08*	0.054	1.09**	0.050	1.39	0.176
Percent of Total Time								
Percent Chart Review	1.20	0.883	1.76	0.698	0.91	0.952	1.96	0.915
Percent Documentation	0.53	0.430	0.25	0.108	0.91	0.944	0.02	0.612
Percent Orders	521.51***	0.003	282.05***	0.010	76.76	0.159	2.83E-23*	0.087
Percent Messaging	0.03	0.217	0.32	0.754	0.13	0.425	37.99	0.898
Percent Problems	1.57E+09*	0.052	5.26E+08*	0.053	0.00	0.791	0.00	0.127
Percent Other	0.67	0.889	5.28E+06	0.100	2.19	0.797	4.17E+03	0.150
Percent After Hours	1.02	0.968	4.30*	0.094	1.77	0.477	4.05	0.575
Computerized Physician Order Entry								
Total Orders Per Patient	1.12**	0.032	1.17*	0.068	1.09	0.120	1.17	0.783
Teamwork for Orders	0.59	0.238	0.73	0.507	0.89	0.862	0.37	0.797
Electronic Documentation Adoption	2.53**	0.012	1.87	0.115	2.10	0.299	3.13E-05	0.588
Adoption Percent	3.72*	0.094	4.15*	0.081	2.92	0.331	7.36	0.797
Decision Support Interaction								
Total Alerts Fired Per Patient	1.01	0.966	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]
Percent Overriden	1.41	0.284	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]
Mobility								
Charts Opened Per Patient	1.50	0.257	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]
Advanced Features								
Percent Dynamic Documentation	0.64	0.421	0.96	0.937	1.04	0.978	NA[2]	NA[2]
Auto-text Per Patient	0.97	0.568	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]
Venue								
Percent Ambulatory	2.40***	0.008	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]
Percent Inpatient	0.48*	0.053	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]	NA[3]

[1] Emergency venue was restricted to only emergency physicians. Inpatient and Ambulatory is restricted to > 25% of work in that venue.

[2] Excluded from analysis due to no variation among emergency physicians.

[3] Excluded from analysis because variable is not measured at the venue level.

Table 6. 2: Multivariate regressions of individual EMR usage variables controlling for gender, age, and specialty

	Measure		Gender-Male		Age 40-54		Age 55+		Non-Surgical Medical		Surgical		Emergency	
Measure	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
Total time in EMR Per Patient	1.03**	0.030	0.64*	0.089	1.93**	0.019	0.85	0.676	0.49**	0.016	1.45	0.351	0.53	0.239
Chart Review Time Per Patient	1.10**	0.026	0.61*	0.060	2.01**	0.013	0.88	0.757	0.47**	0.012	1.41	0.383	0.52	0.219
Documentation Time Per Patient	1.03	0.276	0.62*	0.073	1.84**	0.027	0.79	0.561	0.46***	0.008	1.15	0.706	0.44	0.121
Prescription Time Per Patient	1.40***	0.003	0.59**	0.044	1.94**	0.019	0.90	0.795	0.59*	0.089	1.86	0.139	0.62	0.376
Message Center Time Per Patient	1.45*	0.092	0.61*	0.054	1.79**	0.036	0.77	0.511	0.57*	0.067	1.53	0.314	0.56	0.290
Problems Time Per Patient	5.35**	0.038	0.62*	0.063	1.79**	0.036	0.75	0.476	0.59*	0.096	1.51	0.310	0.55	0.269
Other Time Per Patient	1.65	0.126	0.60*	0.051	1.90**	0.020	0.83	0.643	0.48**	0.012	1.20	0.630	0.26**	0.026
Information Seeking														
Summary (MPage) CR Time Per Patient	0.99	0.947	0.59**	0.041	1.81**	0.031	0.79	0.546	0.47**	0.013	1.10	0.818	0.39*	0.084
Flowsheet Cer Time Per Patient	1.23*	0.075	0.59**	0.039	1.96**	0.017	0.91	0.823	0.43***	0.005	1.20	0.626	0.39*	0.069
Clinical Notes CR Time Per Patient	1.26**	0.011	0.60**	0.048	1.96**	0.016	0.84	0.667	0.42***	0.004	1.26	0.548	0.51	0.205
Other CR Time Per Patient	1.76*	0.052	0.62*	0.068	1.93**	0.019	0.86	0.715	0.47**	0.010	1.29	0.509	0.48	0.171
Work Outside of Work Per Patient	1.09***	0.008	0.59**	0.044	1.89**	0.022	0.85	0.685	0.44***	0.006	1.23	0.584	0.35**	0.047
Percent of Total Time														
Percent Chart Review	3.82	0.350	0.57**	0.030	1.90**	0.022	0.84	0.663	0.46***	0.008	1.12	0.764	0.40*	0.081
Percent Documentation	0.34	0.283	0.56**	0.027	1.84**	0.027	0.81	0.592	0.51**	0.024	1.28	0.534	0.38*	0.062
Percent Orders	1.34E+03***	0.005	0.56**	0.027	1.88**	0.024	0.85	0.681	0.62	0.118	1.84	0.147	0.50	0.200
Percent Messaging	4.78E-04*	0.050	0.61*	0.054	1.98**	0.015	0.91	0.807	0.42***	0.004	0.95	0.888	0.32**	0.033
Percent Problems	6.27E+08	0.137	0.59**	0.039	1.76**	0.042	0.73	0.438	0.56*	0.060	1.29	0.522	0.47	0.159
Percent Other	108.72	0.277	0.59**	0.039	1.86**	0.024	0.79	0.550	0.47***	0.010	1.12	0.755	0.24**	0.047
Percent After Hours	1.51	0.541	0.58**	0.037	1.81**	0.030	0.81	0.592	0.46***	0.009	1.11	0.788	0.35*	0.062
Computerized Physician Order Entry														
Total Orders Per Patient	1.14**	0.019	0.58**	0.036	1.97**	0.016	0.82	0.617	0.51**	0.022	1.41	0.382	0.49	0.182
Teamwork for Orders	0.29*	0.061	0.61*	0.056	1.95**	0.018	0.90	0.788	0.48**	0.013	1.76	0.208	0.51	0.212
Electronic Documentation Adoption	2.55**	0.028	0.60*	0.052	1.91**	0.020	0.81	0.594	0.51**	0.025	1.21	0.614	0.56	0.286
Adoption Percent	5.30	0.125	0.63*	0.076	1.92**	0.019	0.84	0.665	0.49**	0.016	1.51	0.338	0.47	0.150
Decision Support Interaction														
Total Alerts Fired Per Patient	1.14	0.460	0.58**	0.035	1.88**	0.024	0.81	0.608	0.47**	0.010	1.13	0.742	0.41*	0.090
Percent Overriden	1.23	0.568	0.59**	0.042	1.80**	0.032	0.80	0.574	0.48**	0.011	1.13	0.755	0.41*	0.092
Mobility														
Charts Opened Per Patient	1.27	0.514	0.59**	0.041	1.80**	0.032	0.79	0.549	0.47***	0.010	1.09	0.822	0.41*	0.085
Advanced Features														
Percent Dynamic Documentation	0.21**	0.021	0.54**	0.019	1.76**	0.041	0.76	0.487	0.40***	0.003	1.04	0.910	0.42*	0.096
Auto-text Per Patient	0.96	0.545	0.59**	0.039	1.82**	0.030	0.79	0.550	0.47**	0.011	1.09	0.816	0.38*	0.067

OR=1.4, p=.003; OR=.00134, p=.005 respectively). When controlling for demographics, chart review is significantly related to burnout (OR=1.1, p=.026) and different information seeking strategies were more significant in distinguishing burnout (OR=1.26, p=.011 for clinical notes and OR=1.23, p=.075 for flowsheet time per patient) while time spent in summary pages still did not achieve significance.

6.4 Discussion

In the May 2021 special burnout edition of the Journal of the American Medical Informatics Association, Dymek et al proposed a framework to reduce documentation burden using informatics to address clinical note creation, chart review, and inbox management. They base this focus on a heuristic proposed by Arndt et al. of the most burdensome EMR tasks (Arndt et al., 2017). Our evidence here suggests that if we are to reduce burnout, it may require focus on reducing burden of CPOE in terms of total orders, total time spent ordering, and orders as a percent of documentation time. This is consistent with Shanafelt et al.'s 2016 finding that CPOE contributes to burnout. (T. D. Shanafelt et al., 2016) and Overhage and McCallie's 2020 evidence from one large national study that CPOE consumes 17% of total EMR time, behind only chart review and documentation. One pre-post study showed that order entry clerical staff reduced burnout in one practice from 3 physicians to 1 (of 7), but we know of no studies that have scaled up this type of analysis. Additionally, more orders burden may tie to higher acuity patients, a confounder that we could not control for in this analysis, but which may be driving burnout more than the measurable time spent on CPOE. Conversely, although notes documentation has been demonstrated to take up 24% of physician EMR time, we found no evidence that this is contributing to physician burnout. Like other studies, evidence of the impact of chart review time was mixed (Domaney et al., 2018; Hilliard et al., 2020), showing a significant contribution to burnout only when we controlled for individual level demographic factors, suggesting a possible need for focus on women aged 40-54 in primary care. Inbox management is another area that has gotten a lot of attention. Our findings were also mixed in this space, uncovering evidence of a relationship between message center time across venues and in the ambulatory setting that attenuated when we controlled for demographic characteristics. We hypothesize that this is in part due to the variety of tasks such as patient call messages, results messages, care team communication, and that the impact of inbox burden is exacerbated when there is not sufficient time to handle the volume (Gregory et al., 2017).

One area that does warrant attention is problem list management. This activity accounts for a small proportion of the time in EMR. In fact, this variable was almost excluded from the analysis as 31% of the included physicians recorded no time managing the problem list. However, problems time was highly significant at the .01 level (OR=5.78) suggesting that additional problem list burden would contribute greatly to burnout. However, we know that the problem list can potentially drive intelligence in the EMR. For example, one intervention demonstrated to reduce cognitive load and burnout leverages a problem-oriented clinical data retrieval strategy (Semanik et al., 2021), but such informatics-based improvements are not possible without a well-managed problem list, something that we anecdotally know many physicians have been slow to adopt.

Previous exploratory data analysis had suggested different patterns of EMR usage across venues. We extend that finding here with the finding that EMR usage more highly correlates with burnout in the ambulatory setting but has almost no relationship in the inpatient and emergency. This could be for several reasons including in-venue variation as well as the problematic nature of assigning a physician to a particular venue. For example, a surgeon may spend a day and a half in surgery seeing 5 total inpatient surgical cases in a week but may see 25 patients in the ambulatory setting as consultations or follow-up. This provider would be attributed as an ambulatory physician, but like a more nuanced understanding of their time in each venue is needed.

Not terribly surprising there were no EMR usage variables that positively affect burnout (OR<1). Work outside of work probably warrants some discussion. Based on how this was operationalized by the vendor (6:00am to 6:00pm), this conceptually more closely represents time worked outside of core hours in the ambulatory venue and shift work in the inpatient and emergency venues. Interestingly, this approached significance in the ambulatory setting (p=.054, OR=1.08) and in the inpatient setting (p=.05, OR=.054) suggesting that both documentation time outside of core hours and shift work may be related to burnout.

Individual-level characteristics unearthed some interesting relationships to burnout. Not terribly surprising, male gender decreased odds of burnout compared with women by about 50%. Age shows a non-linear relationship to burnout with those in the middle age range from 40-45 significantly having twice the odds of burnout compared with those physicians under the age of 40. However, no significant difference in burnout was found between the < 40 group and the > 55 group. Finally, Non-surgical medical specialties had about half the odds of burnout of primary care which minimal significant difference was found between primary care and surgical and emergency specialties.

There was some limited evidence of decreased odds of burnout in emergency medicine compared to primary care, but we caution against over-interpreting this due to low sample size.

Overall, this analysis demonstrated the promise of an EMR-based system to support a learning health system approach to improving EMR contributions to physician burnout. The results here reinforce findings of other studies regarding the relationship between both individual physician characteristics and EMR usage to burnout. It highlights some potential areas for focus such as CPOE and problem list management. The analysis was intentionally done across venues and to allow for sub-group analysis to see how the approach might scale. In a true LHS, the operationalization of event log data would need to adapt to more accurately reflect the specific EMR behavior at which the improvement is targeted. For example, in the inpatient setting nurses and physicians in a traditional EMR may repeatedly refresh the results view waiting for critical lab results to return. An intervention focused on improving this, for example, an intelligent mobile based notification system, may reduce overall chart review time, but more specifically the impact may be better measured through a reduction in repeated refresh of the results tab while waiting for a critical result.

6.5 Limitations

This study suffers from several limitations. First, the pre-aggregated data do not allow for a more granular characterization of EMR usage. This study is intended to demonstrate the feasibility and utility of an EMR-based system to understand EMR contributions to burnout, and thus any interventions to improve the EMR are out of scope – although we suggest that this is a very active area of research and will continue to be a focus in the future. Further, while the pre-aggregated data have been validated elsewhere, as several authors have called out, these definitions are not consistent across EMR vendors, which limits their utility. Some variables would benefit from enhancement. For example, the per patient denominator would be better characterized by the number of patient visits since we know that physicians don't always complete notes at the same time as they finalize notes – although we believe this affect is mitigated by the aggregation at the monthly level. Similarly, to truly operationalize work outside of work, it is necessary to compare EMR time with scheduling data, which was out of scope for the current study. The characterization of venue is problematic given the extent to which physicians move between venues, and, in particular, between the inpatient and ambulatory venues, but we believe this is an important distinction and find that varying the threshold at which a physician is considered "ambulatory" or "inpatient" did not meaningfully impact the results. This was a single site study and, as such, results cannot be generalized. Finally, many patient, provider, and organizational level confounders were not able to be accounted for in this study and certainly no causal connections can be drawn from any of the analyses.

6.6 Conclusion:

This study demonstrates the utility of an EMR-based model for repeated measure of EMR contributions to burnout and scale-up of the measurement of impact of interventions within a Learning Health System. Our results reinforce existing evidence on the relationship of individual-level characteristics and EMR usage with burnout. To apply this system to pre-post intervention measurement of interventions to reduce EMR causes of burnout, future work should focus on enhancing how we use event log data to measure relevant EMR activity and scale up, possibly across sites, in a randomized fashion to demonstrate causal impact of interventions to reduce EMR-related physician burnout.

Chapter 7: Conclusion

A national and international conversation on how to reduce documentation burden and the associated physician burnout is well underway. In January and February of 2021, over 300 participants from 140 organizations attended a series of weekly sessions to outline the key issues and opportunities to address this epidemic (Rossetti et al.). Shortly after, the May 2021 special burnout edition of the Journal of the American Medical Informatics Association (JAMIA) focused entirely on this issue synthesizing evidence around the relationship of Health Information Technology (HIT) to clinician burnout, highlighting emerging approaches to mitigate burnout, and providing perspectives for future directions to combat this trend (Poon, Trent Rosenbloom, & Zheng, 2021). This dissertation adds to that very active dialogue by addressing a gap in our ability to easily measure the impact of these efforts

7.1 Dissertation Overview and Findings

This dissertation demonstrated the feasibility, reliability, and utility of a first-of-its-kind EMR-based system for measuring and correlating EMR use with physician burnout. The literature review in Chapter 3 uncovered a large body of evidence linking EMR use to physician well-being. While there were some mixed results directly connecting EMR use to burnout, we uncovered clear evidence that the confounding factor of limited time, or a perception of limited time, to address documentation did contribute to burnout. While this body of literature is growing in the last couple of years, it remains highly observational with only three pre-post intervention studies uncovered in the review. We also uncovered a trend away from the gold standard Maslach Burnout Inventory in preference of simpler instruments such as the Mini-Z or one of the validated

Single-Item Burnout Measures. We also uncovered an increase in use of secondary event log data to characterize different aspects of EMR use as well as efforts underway to develop frameworks for defining key measures of EMR usage and to standardize across EMR vendors. Within this backdrop, we proposed to work toward closing the gap in prospective studies by validating an EMR-based measurement system for relating EMR use to burnout.

7.1.1 Approach

We demonstrated the ability to correlate EMR usage to burnout in several steps: feasibility of collecting repeated measure of burnout in the EMR, data mining to characterize clinical workflow patterns through use of existing EMR event log data, and finally identified correlative relationships between EMR use and burnout. We first piloted an externally validated SIBM sent through the EMR message center. Adjustments were made to our privacy protocols and communication plan based on feedback, and the survey was released once a month for three months to cohorts that were randomized by time of day and day of week. Chi-square, Mann-Whitney U test, and Kruskal Wallace tests were performed to identify how individual-level characteristics affect response rates and burnout scores. We further tested to ensure that randomized survey administration did not impact response rates and burnout scores.

To characterize EMR usage patterns, we obtained pre-processed data aggregated by physician monthly usage. Data included time in EMR, including information-seeking patterns, documentation outside of core work hours, Computerized Physician Order Entry (CPOE) usage, electronic adoption percentages, EMR performance and stability, Decision support burden and interaction, mobility, and use of advanced features. For each physician, we tested consistency of EMR usage month over month within a venue and between the ambulatory and inpatient venue for physicians who practice in both. We further sought to identify latent constructs of EMR through use of factor analysis and further thought to see if these latent constructs varied for physicians practicing within a single venue (inpatient or ambulatory). Finally, we used logistic regression to check for correlation between the data as characterized and the burnout scores as captured through our SIBM measure in the workflow.

7.1.2 Major Findings

We uncovered significant biases in overall response rates by gender and specialty with women and primary care physicians responding at a much higher rate than their counterparts. As hypothesized, the randomized survey administration by day of week and time of day did not have a significant impact on response rates. The month of survey administration did have a drop off in response rate from month 1 to 2 and month 2 to 3. We further discovered, consistent with the literature, that individual characteristics of gender, age, and specialty were significantly related to burnout scores at a 99% significance level. Again, none of the survey administration or response variables was significantly related to burnout. There is some evidence that of higher burnout for physicians surveyed on Fridays, but it failed to reach a 95% significance threshold in a Kruskal-Wallis test of significance.

We found good to excellent consistency, as measured by intra-class correlation (ICC) for most EMR usage variables over time. The exceptions were use of mobility, decision support behavior, and use of the advanced autotext feature. This high ICC did not hold true when looking at physician practice between the ambulatory and inpatient venues, with ICC failing to reach a good level of .75 at 95% significance for any variables. Our exploratory factor analysis failed to unearth clearly identifiable latent classes. Five to six factors were needed to reach a generally accepted Tucker Lewis Index (TLI) of .9. Furthermore, the models only explained between .45 and .71 of the cumulative variance. Logistic regression showed that the factors in the overall model appear to be a tradeoff between the inpatient and ambulatory venue, with logistic regression showing a significant relationship with ambulatory physicians, and, for factors 2-4 which were also significantly related to an inpatient flag, showing an inverse relationship to inpatient physicians. The ambulatory and inpatient models showed similar challenges with cross-loading and identification of meaningful latent patterns. However, logistic regression showed a strong relationship between EMR use patterns and specialty in the ambulatory venue. This relationship largely did not hold true for the inpatient venue.

Finally, our univariate regression analyses uncovered many highly significant relationships between EMR usage and burnout – specifically total time in EMR, orders, message center, and problems time, and, to a lesser extent, for chart review. The percent of ambulatory time increased the odds of burnout, while, converse, the percent of time in inpatient significantly decreased the odds of burnout ($p=.053$). The patterns within the ambulatory venue almost exactly match the relationships uncovered overall. Virtually no evidence was found of relationship between EMR and burnout in the inpatient and emergency venues except for work outside of the hours of 6:00 am and 6:00 pm which significantly increased the odds of burnout. These patterns remained consistent when controlling for gender, age, and specialty with the noteworthy exception that chart review, and associated information seeking patterns, reached significance within the multivariate model. The multivariate model also revealed a highly significant gender relationship with men showing around half the odds of burnout as women. Physicians aged 40-54 had around 2 times the odds of burnout compared to physicians under 40, but there was no significant change in odds of burnout between the < 40 group and the > 55 group. Non-surgical medical specialties had about half the odds of burnout of primary care which virtually no significant difference was found between primary care and surgical and emergency specialties.

7.2 Discussion and Future Directions

This paper provides evidence of the feasibility and reliability of using a SIBM deployed in the clinical workflow to capture repeated measures of burnout which can potentially facilitate future prospective examination of impact of interventions. We were further able to demonstrate how these burnout scores, coupled with EMR event logging data can be used to correlate EMR usage behavior to burnout. While our use of existing aggregate EMR use measures proved somewhat problematic in exploratory analysis to identify latent clinical workflows, we did uncover evidence that differentiates between venue EMR usage as well as differences between specialties. We were further able to provide evidence using this EMR-based measurement system that is consistent with existing findings in the literature on how the EMR relates to burnout. As mentioned in the introduction, this research was undertaken with future steps in mind, and we believe our findings advise several future directions.

First, where improvement is underway, we can leverage the SIBM to test the utility of this system for measuring improvement efforts already planned or in flight.

Future research needs to test the utility of the SIBM in the workflow to see if sensitivity and specificity are adequate to use as a pre-post intervention measurement. We suspect that the SIBM will be one additional tool for prospectively measuring impact, but a critical one to demonstrate the industry as a whole is making headway in combatting this pandemic. Second, improvement in how we use event log data to understand clinical workflows is already under way. As mentioned previously, several researchers are working to standardize metrics of EMR usage across EMR vendors. However, we agree with Adler-Milstein, Adelman, Tai-Seale, Patel, & Dymek (2020) that EMR event log data provide greater opportunities across multiple quality domains for health services research. We also look to informatics and industrial engineering, and computer science colleagues to continue to advance methods for using those data to clinical workflow as a network of caregivers or sequence of activities. Finally, with advancements in methods to understand clinical workflow through event logging data coupled with a reliable SIBM measure deployed in the EMR for repeated measures, we anticipate that in the near future predictive systems are possible which can predict future EMR-related burnout and mitigate it before it occurs.

APPENDICES

Appendix A: Single Item Burnout Measures Cited

Rohland et al. 2004 Single Item Burnout Measure

Overall, based on your definition of burnout, how would you rate your level of burnout?

1) I enjoy my work, I have no symptoms of burnout
2) Occasionally I am under stress, and I don't always have as much energy as I once did, but I don't feel burned out
3) I am definitely burning out and I have one or more symptoms of burnout, such as physical and emotional exhaustion
4) The symptoms of burnout that I'm experiencing won't go away. I think about frustration at work a lot
5) I feel completely burned out and often wonder if I can go on. I am at the point where I may need some changes or may need some help.

Mini-Z Single-Item Burnout Measure

Using your own definition of burnout, please circle one of the answers below:

a) I enjoy my work. I have no symptoms of burnout.
b) I am under stress, and don't always have as much energy as I did, but I don't feel burned out.
c) I am definitely burning out and have one or more symptoms of burnout, e.g., emotional exhaustion.
d) The symptoms of burnout that I am experiencing won't go away. I think about work frustrations a lot
e) I feel completely burned out. I am at the point where I may need to seek help.

Gilleland et al. 2014 Single Question fatigue/burnout survey

(1) "I am serenity, life is fine."
(2) "I have some stress, but things are good."
(3) "I am moderately stressed out but usually handle it well."
(4) "I am pretty stressed out and sometimes it affects my performance."
(5) "I am burned out."

Maslach et al. 1986: Maslach Burnout Inventory single item questions referenced

Emotional Exhaustion: I feel burned out from my work
Depersonalization: I have become more callous toward people since I took this job.

Never
A few times a year or less
Once a month or less
A few times a month
Once a week
A few times a week
Every Day

Appendix B: Audit log variables

Measure	Description	Standardization	Available by Venue
Total time in EMR	Total hours per spent spent in the EMR during and outside of the clinical visit	Per Patient	Y
Chart Review Time	Hours per patient spent reviewing the flowsheets, summary pages, clinical notes and "other" time.	Per Patient	Y
Documentation Time	Hours spent per patient creating the physician clinical note	Per Patient	Y
Prescription Time	Hours per patient spent on computerized physician order entry.	Per Patient	Y
Message Center Time	Hours spent in the EMR inbox messaging members of the care team, external providers, or the patient	Per Patient	Y
Problems Time	Hours spent entering problems into the problem list	Per Patient	Y
Other Time	Bucket for all other time	Per Patient	Y
Information Seeking			
Summary (MPage)	Chart review time spent in summary pages (called MPages)	Per Patient	Y
Flowsheet	Chart review time spent in the flowsheet reviewing vitals and results	Per Patient	Y
Clinical Notes	Chart review time spent reviewing clinical notes	Per Patient	Y
Other	Other Chart Review time	Per Patient	Y
Work Outside of Work	Measure of time spent in the EMR outside of core work hours defined as 6:00am to 6:00 pm	Per patient	Y
Computerized Physician Order Entry			
Total Orders	Electronic Orders signed by the phsyician	Per Patient	Y
Teamwork for Orders	Percentage of orders signed by the originating physician. Designed to capture order that originate from residents or protocol	% of total orders	Y
Electronic Documentation Adoption	Percentage of notes that originated in the EMR and were electronically signed by the physician	% of total notes	Y
Decision Support Interaction			
Total Alerts Fired	Total number of decision support alerts experienced by physician	Per Patient	N
Percent Overriden	Percent of overridable alerts that were overridden by the physician	% of total alerts	N
Mobility	Captures the extent to which documentation happens on a mobile (iPad, Surface, smart phone)		
Charts Opened	Total charts opened in mobile application	Per Patient	N
Advanced Features			
Dynamic Documentation	Percent of clinical notes using advanced dynamic documentation	% of total notes	Y
Auto-text	Total number of times a physician uses pre-defined note templates	Per Patient	N
Venue			
Percent Ambulatory	Percent of total patients seen in an ambulatory venue	Total Patients	N
Percent Inpatient	Percent of total patients seen in an inpatient venue	Total Patients	N
Percent Emergency	Percent of total patients seen in the emergency department	Total Patients	N

Appendix C: Full Factor Analysis Results

```
Factor Analysis using method =  minres
Call: fa(r = Numeric_FADF, nfactors = 6, rotate = "oblimin",
    max.iter = 100, fm = "minres")
Standardized loadings (pattern matrix) based upon correlation matrix
                        MR3   MR2   MR4   MR5   MR6   MR1    h2     u2 com
DocTimePP             -0.02  0.21  0.34  0.54 -0.16 -0.05 0.580 0.4204 2.3
OrderTimePP            0.75  0.16  0.07  0.11  0.08 -0.07 0.862 0.1379 1.2
MessagingTimePP        0.19  0.36 -0.06  0.27  0.19 -0.12 0.479 0.5212 3.4
PatDiscoveryTimePP     0.13  0.19  0.05  0.21  0.21  0.64 0.975 0.0252 1.8
ProblemsTimePP         0.23  0.59  0.02 -0.02  0.04 -0.22 0.549 0.4512 1.6
OtherTimePP           -0.04  0.27  0.15 -0.18  0.23  0.22 0.243 0.7570 4.4
ActualTimeAfterHoursPP  0.02  0.00  0.97  0.01  0.04 -0.01 0.996 0.0038 1.0
ordersPP               0.87 -0.04  0.00 -0.04 -0.07  0.10 0.726 0.2745 1.0
TWFOrders             -0.50 -0.09  0.10 -0.07 -0.15  0.19 0.347 0.6532 1.7
TotalAlertsFiredPP     0.45 -0.11  0.23 -0.10 -0.02  0.33 0.469 0.5312 2.7
OverridePercentage     0.23 -0.02  0.04 -0.12  0.24 -0.13 0.108 0.8917 3.1
MobileChartsOpenedPP   0.05 -0.01  0.04 -0.09  0.30  0.17 0.161 0.8392 1.9
MPageCRTimePP         -0.01  0.91  0.01 -0.01  0.00  0.12 0.842 0.1583 1.0
FlowsheetCRTimePP      0.09 -0.12  0.29  0.09  0.53  0.17 0.727 0.2733 2.1
NotesCRTimePP          0.05 -0.09  0.00  0.75  0.15  0.12 0.726 0.2743 1.2
OtherCRTimePP         -0.05  0.13  0.06  0.11  0.61  0.03 0.560 0.4404 1.2
AutoTextPP             0.01  0.12 -0.02  0.14 -0.02  0.04 0.044 0.9557 2.2

                       MR3  MR2  MR4  MR5  MR6  MR1
SS loadings           2.25 1.81 1.55 1.36 1.41 1.01
Proportion Var        0.13 0.11 0.09 0.08 0.08 0.06
Cumulative Var        0.13 0.24 0.33 0.41 0.49 0.55
Proportion Explained  0.24 0.19 0.17 0.14 0.15 0.11
Cumulative Proportion 0.24 0.43 0.60 0.74 0.89 1.00

 With factor correlations of
     MR3  MR2  MR4  MR5  MR6  MR1
MR3 1.00 0.41 0.41 0.28 0.35 0.22
MR2 0.41 1.00 0.20 0.37 0.28 0.04
MR4 0.41 0.20 1.00 0.44 0.40 0.36
MR5 0.28 0.37 0.44 1.00 0.47 0.22
MR6 0.35 0.28 0.40 0.47 1.00 0.41
MR1 0.22 0.04 0.36 0.22 0.41 1.00

Mean item complexity =  2
Test of the hypothesis that 6 factors are sufficient.

The degrees of freedom for the null model are  136  and the objective function was  7.52 with Chi Square of  10409.08
The degrees of freedom for the model are 49  and the objective function was  0.28

The root mean square of the residuals (RMSR) is  0.02
The df corrected root mean square of the residuals is  0.03

The harmonic number of observations is  1392 with the empirical chi square  140.69  with prob <  8.5e-11
The total number of observations was  1392  with Likelihood Chi Square =  384.54  with prob <  1.3e-53

Tucker Lewis Index of factoring reliability =  0.909
RMSEA index =  0.07  and the 90 % confidence intervals are  0.064 0.077
BIC =  29.86
Fit based upon off diagonal values = 1
Measures of factor score adequacy
                                                  MR3  MR2  MR4  MR5  MR6  MR1
Correlation of (regression) scores with factors   0.95 0.94 1.00 0.89 0.87 0.96
Multiple R square of scores with factors          0.89 0.88 1.00 0.79 0.76 0.92
Minimum correlation of possible factor scores     0.79 0.76 0.99 0.57 0.52 0.83
```

C.1: Results for all venue for two-way random effects for consistency with single-rater factor analysis after feature reduction using OLS method with oblique rotation.

```
Factor Analysis using method =  minres
Call: fa(r = Numeric_FADF, nfactors = 6, rotate = "oblimin",
    max.iter = 100, fm = "minres")
Standardized loadings (pattern matrix) based upon correlation matrix
                        MR2   MR1   MR5   MR3   MR6   MR4   h2     u2 com
DocTimePP              0.14 -0.08  0.65  0.05  0.06  0.29 0.67 0.3291 1.5
OrderTimePP            0.06  0.72  0.11 -0.01  0.18  0.07 0.83 0.1697 1.2
MessagingTimePP        0.23  0.43  0.13  0.23  0.12 -0.27 0.67 0.3254 3.4
PatDiscoveryTimePP     0.96  0.01 -0.04  0.14  0.00  0.03 1.00 0.0025 1.0
ProblemsTimePP        -0.10  0.52  0.07  0.40 -0.07  0.18 0.57 0.4308 2.3
OtherTimePP            0.36  0.15 -0.14  0.36 -0.05  0.08 0.36 0.6394 2.8
ActualTimeAfterHoursPP -0.05  0.12  0.77  0.09 -0.03 -0.11 0.65 0.3477 1.1
ordersPP               0.07  0.49 -0.11 -0.01  0.37  0.02 0.56 0.4416 2.0
TWFOrders              0.05 -0.03  0.02 -0.03 -0.83 -0.06 0.73 0.2673 1.0
ElectDocPercentAuthored 0.07  0.11  0.04  0.03  0.15  0.67 0.62 0.3778 1.2
MPageCRTimePP          0.13  0.02  0.12  0.78  0.11  0.03 0.85 0.1547 1.2
FlowsheetCRTimePP      0.41  0.36  0.19 -0.18  0.05 -0.20 0.58 0.4173 3.4
NotesCRTimePP          0.63  0.03  0.30 -0.25  0.04  0.04 0.68 0.3158 1.8
OtherCRTimePP          0.22 -0.09  0.15  0.24  0.31 -0.14 0.32 0.6785 4.0

                       MR2  MR1  MR5  MR3  MR6  MR4
SS loadings           2.19 1.88 1.54 1.35 1.35 0.79
Proportion Var        0.16 0.13 0.11 0.10 0.10 0.06
Cumulative Var        0.16 0.29 0.40 0.50 0.59 0.65
Proportion Explained  0.24 0.21 0.17 0.15 0.15 0.09
Cumulative Proportion 0.24 0.45 0.62 0.77 0.91 1.00

 With factor correlations of 
     MR2  MR1  MR5  MR3  MR6  MR4
MR2 1.00 0.38 0.52 0.29 0.36 0.04
MR1 0.38 1.00 0.37 0.34 0.55 0.02
MR5 0.52 0.37 1.00 0.16 0.25 0.09
MR3 0.29 0.34 0.16 1.00 0.24 0.21
MR6 0.36 0.55 0.25 0.24 1.00 0.30
MR4 0.04 0.02 0.09 0.21 0.30 1.00

Mean item complexity =  2
Test of the hypothesis that 6 factors are sufficient.

The degrees of freedom for the null model are  91  and the objective function was  7.08 with Chi Square of  7538.78
The degrees of freedom for the model are 22  and the objective function was  0.17 

The root mean square of the residuals (RMSR) is  0.01 
The df corrected root mean square of the residuals is  0.03 

The harmonic number of observations is  1072 with the empirical chi square  42.41  with prob <  0.0056 
The total number of observations was  1072  with Likelihood Chi Square =  185.2  with prob <  8.7e-28 

Tucker Lewis Index of factoring reliability =  0.909
RMSEA index =  0.083  and the 90 % confidence intervals are  0.072 0.095
BIC =  31.7
Fit based upon off diagonal values = 1
Measures of factor score adequacy             
                                                   MR2  MR1  MR5  MR3  MR6  MR4
Correlation of (regression) scores with factors   1.00 0.92 0.90 0.92 0.89 0.83
Multiple R square of scores with factors          0.99 0.85 0.80 0.84 0.80 0.68
Minimum correlation of possible factor scores     0.99 0.71 0.61 0.68 0.60 0.37
```

C.2: Results for Ambulatory venue for two-way random effects for consistency with single-rater factor analysis after feature reduction using OLS method with oblique rotation.

```
Factor Analysis using method =  minres
Call: fa(r = Numeric_FADF, nfactors = 5, rotate = "oblimin", 
    max.iter = 100, fm = "minres")
Standardized loadings (pattern matrix) based upon correlation matrix
                         MR1   MR2   MR3   MR4   MR5   h2      u2 com
DocTimePP               0.06  0.00  0.94  0.06  0.02 1.00 0.00087 1.0
OrderTimePP             0.05  0.70  0.03  0.19  0.25 0.89 0.11097 1.4
PatDiscoveryTimePP      0.60  0.25 -0.17  0.21  0.24 0.78 0.22316 2.2
ProblemsTimePP         -0.09  0.35  0.16  0.45 -0.08 0.50 0.49841 2.3
OtherTimePP             0.14 -0.15  0.01  0.37  0.32 0.33 0.67043 2.6
ActualTimeAfterHoursPP  0.18  0.18  0.42 -0.03  0.57 0.93 0.06780 2.3
ordersPP               -0.02  0.88 -0.06  0.00  0.02 0.73 0.26879 1.0
TWFOrders              -0.17 -0.47 -0.04 -0.04  0.12 0.31 0.69306 1.4
ElectDocPercentAuthored  0.02  0.66  0.21  0.04 -0.19 0.57 0.42527 1.4
MPageCRTimePP           0.07  0.10  0.09  0.72 -0.04 0.70 0.29507 1.1
FlowsheetCRTimePP       0.79 -0.09  0.07  0.06  0.19 0.87 0.12835 1.2
NotesCRTimePP           0.92  0.10  0.09 -0.12 -0.08 0.86 0.13883 1.1
OtherCRTimePP           0.77 -0.13  0.06  0.19 -0.09 0.69 0.30891 1.2

                       MR1  MR2  MR3  MR4  MR5
SS loadings           2.95 2.50 1.49 1.38 0.86
Proportion Var        0.23 0.19 0.11 0.11 0.07
Cumulative Var        0.23 0.42 0.53 0.64 0.71
Proportion Explained  0.32 0.27 0.16 0.15 0.09
Cumulative Proportion 0.32 0.59 0.76 0.91 1.00

 With factor correlations of 
     MR1  MR2  MR3  MR4  MR5
MR1 1.00 0.28 0.60 0.45 0.44
MR2 0.28 1.00 0.34 0.49 0.24
MR3 0.60 0.34 1.00 0.32 0.22
MR4 0.45 0.49 0.32 1.00 0.24
MR5 0.44 0.24 0.22 0.24 1.00

Mean item complexity =  1.6
Test of the hypothesis that 5 factors are sufficient.

The degrees of freedom for the null model are  78  and the objective function was  9.45 with Chi Square of  5424.47
The degrees of freedom for the model are 23  and the objective function was  0.3 

The root mean square of the residuals (RMSR) is  0.02 
The df corrected root mean square of the residuals is  0.03 

The harmonic number of observations is  580 with the empirical chi square  24.72  with prob <  0.36 
The total number of observations was  580  with Likelihood Chi Square =  170.96  with prob <  1.4e-24 

Tucker Lewis Index of factoring reliability =  0.906
RMSEA index =  0.105  and the 90 % confidence intervals are  0.091 0.12
BIC =  24.61
Fit based upon off diagonal values = 1
```

C.3: Results for Inpatient venue for two-way random effects for consistency with single-rater factor analysis after feature reduction using OLS method with oblique rotation.

Appendix D: Specialty Groupings

Specialty Group	**Specialty**
Primary Care	Family Medicine, General Internal Medicine, Pediatrics, Student Health Attending, Women's Health
Surgical Specialties	Acute Care Surgery, General Surgery, Ophthalmology, Orthopaedic, Otolaryngology, Plastic Surgery, Urology
Non-Surgical Medical	Behavioral Health, Cardiology, Dermatology, Endocrinology, Gastroenterology, Internal Medicine (excluding GIM), Nephrology, Neurology, Oncology, Optometrist, PM&R, Pulmonology
Emergency	Emergency

Appendix E: Detailed Intra-Class Correlation Results for all models

Metric	Valuable	Subjects	Raters	F-Range	ICC	F	P=Value	95th Lower	95th Upper
Ambulatory to Inpatient	TWFOrders	300	2	F(299,299)	0.699	5.65	0	0.637	0.753
Ambulatory to Inpatient	NotesCRTimePP	308	2	F(307,307)	0.628	4.38	0	0.555	0.691
Ambulatory to Inpatient	PatDiscoveryTimePP	308	2	F(307,307)	0.62	4.27	0	0.546	0.684
Ambulatory to Inpatient	MessagingTimePP	308	2	F(307,307)	0.616	4.21	0	0.542	0.681
Ambulatory to Inpatient	CRTimePP	308	2	F(307,307)	0.588	3.86	0	0.51	0.657
Ambulatory to Inpatient	DDPercent	308	2	F(307,307)	0.585	3.82	0	0.506	0.654
Ambulatory to Inpatient	FlowsheetCRTimePP	308	2	F(307,307)	0.565	3.6	0	0.484	0.636
Ambulatory to Inpatient	MPageCRTimePP	308	2	F(307,307)	0.542	3.36	0	0.458	0.616
Ambulatory to Inpatient	ActualTimePP	308	2	F(307,307)	0.528	3.24	0	0.443	0.604
Ambulatory to Inpatient	DocTimePP	308	2	F(307,307)	0.452	2.65	0	0.358	0.536
Ambulatory to Inpatient	OtherCRTimePP	308	2	F(307,307)	0.424	2.47	0	0.328	0.511
Ambulatory to Inpatient	ActualTimeAfterHoursPP	308	2	F(307,307)	0.419	2.44	0	0.322	0.507
Ambulatory to Inpatient	OtherTimePP	308	2	F(307,307)	0.391	2.29	0	0.293	0.482
Ambulatory to Inpatient	ElectDocPercentAuthorec	305	2	F(304,304)	0.359	2.12	0	2.59E-01	0.453
Ambulatory to Inpatient	ProblemsTimePP	308	2	F(307,307)	0.17	1.41	0.00139	0.059	0.276
Ambulatory to Inpatient	OrderTimePP	308	2	F(307,307)	0.129	1.3	0.0117	0.018	0.237
Ambulatory to Inpatient	OrdersPP	308	2	F(307,307)	0.118	1.27	0.0188	0.007	0.227
Month Over Month Within Venu	DDPercent	531	3	F(530,1060)	0.963	78.1	0	0.957	0.968
Month Over Month Within Venu	OtherTimePP	531	3	F(530,1060)	0.917	34.2	0	0.905	0.928
Month Over Month Within Venu	MPageCRTimePP	531	3	F(530,1060)	0.908	30.7	0	0.895	0.92
Month Over Month Within Venu	TWFOrders	507	3	F(506,1012)	0.907	30.4	0	0.894	0.92
Month Over Month Within Venu	ElectDocPercentAuthorec	528	3	F(527,1054)	0.905	29.5	0	0.891	0.917
Month Over Month Within Venu	DocTimePP	531	3	F(530,1060)	0.903	28.9	0	0.889	0.916
Month Over Month Within Venu	ActualTimePP	531	3	F(530,1060)	0.901	28.3	0	0.886	0.914
Month Over Month Within Venu	ProblemsTimePP	531	3	F(530,1060)	0.893	26	0	0.878	0.907
Month Over Month Within Venu	OrderTimePP	531	3	F(530,1060)	0.877	22.4	0	0.86	0.893
Month Over Month Within Venu	PatDiscoveryTimePP	531	3	F(530,1060)	0.874	21.9	0	0.857	0.891
Month Over Month Within Venu	CRTimePP	531	3	F(530,1060)	0.871	21.2	0	0.852	0.887
Month Over Month Within Venu	OrdersPP	531	3	F(530,1060)	0.871	21.2	0	0.852	0.887
Month Over Month Within Venu	FlowsheetCRTimePP	531	3	F(530,1060)	0.868	20.7	0	0.849	0.885
Month Over Month Within Venu	MessagingTimePP	531	3	F(530,1060)	0.863	19.9	0	0.844	0.881
Month Over Month Within Venu	NotesCRTimePP	531	3	F(530,1060)	0.851	18.1	0	0.83	0.87
Month Over Month Within Venu	ActualTimeAfterHoursPP	531	3	F(530,1060)	0.835	16.1	0	0.812	0.855
Month Over Month Within Venu	OtherCRTimePP	531	3	F(530,1060)	0.8	13	0	0.773	0.824
Month Over Month All Venue	PercED	455	3	F(454,908)	0.987	225	0	0.985	0.989
Month Over Month All Venue	PercAmb	455	3	F(454,908)	0.928	39.7	0	0.917	0.938
Month Over Month All Venue	PercInp	455	3	F(454,908)	0.906	30.1	0	0.892	0.92
Month Over Month All Venue	TotalAlertsFiredPP	455	3	F(454,908)	0.868	20.7	0	0.848	0.886
Month Over Month All Venue	AverageTransactionResp	455	3	F(454,908)	0.759	10.5	0	0.726	0.79
Month Over Month All Venue	MobileChartsOpenedPP	455	3	F(454,908)	0.715	8.54	0	0.677	0.751
Month Over Month All Venue	TotalCrashCountPP	455	3	F(454,908)	0.577	5.1	0	0.528	0.625
Month Over Month All Venue	OverridePercentage	455	3	F(454,908)	0.528	4.36	0	0.476	0.579
Month Over Month All Venue	AutoTextPP	455	3	F(454,908)	0.116	1.39	1.6E-05	0.06	0.175

BIBLIOGRAPHY

Adler-Milstein, J., Adelman, J. S., Tai-Seale, M., Patel, V. L., & Dymek, C. (2020). EHR audit logs: A new goldmine for health services research? *Journal of Biomedical Informatics*, *101*, 103343. doi:https://doi.org/10.1016/j.jbi.2019.103343

Adler-Milstein, J., & Jha, A. K. (2017). HITECH Act Drove Large Gains In Hospital Electronic Health Record Adoption. *Health Aff (Millwood), 36*(8), 1416-1422. doi:10.1377/hlthaff.2016.1651

Adler-Milstein, J., Zhao, W., Willard-Grace, R., Knox, M., & Grumbach, K. (2020). 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. *Journal of the American Medical Informatics Association, 27*(4), 531-538.

Amoafo, E., Hanbali, N., Patel, A., & Singh, P. (2014). What are the significant factors associated with burnout in doctors? *Occupational Medicine, 65*(2), 117-121. doi:10.1093/occmed/kqu144

Arndt, B. G., Beasley, J. W., Watkinson, M. D., Temte, J. L., Tuan, W. J., Sinsky, C. A., & Gilchrist, V. J. (2017). Tethered to the EHR: Primary Care Physician Workload Assessment Using EHR Event Log Data and Time-Motion Observations. *Ann Fam Med, 15*(5), 419-426. doi:10.1370/afm.2121

Astudillo, W., & Mendinueta, C. (1996). Exhaustion syndrome in palliative care. *Support Care Cancer, 4*(6), 408-415. doi:10.1007/bf01880637

Azam, K., Khan, A., & Alam, M. T. (2017). Causes and adverse impact of physician burnout: a systematic review. *J Coll Physicians Surg Pak, 27*(8), 495-501.

Aziz, F., Talhelm, L., Keefer, J., & Krawiec, C. (2019). Vascular surgery residents spend one fifth of their time on electronic health records after duty hours. *J Vasc Surg, 69*(5), 1574-1579. doi:10.1016/j.jvs.2018.08.173

Baxter, S. L., Apathy, N. C., Cross, D. A., Sinsky, C., & Hribar, M. R. (2021). Measures of electronic health record use in outpatient settings across vendors. *J Am Med Inform Assoc*, *28*(5), 955-959. doi:10.1093/jamia/ocaa266

Burisch, M. (1984). Approaches to personality inventory construction: a comparison of merits. *American psychologist*, *39*(3), 214.

Burke, C., Surawicz, C. M., Oxentenko, A. S., Lopez, R., Sarvepalli, S., Anderson, J. C., . . . Richman, S. (2017). A National Survey of Burnout in Gastroenterologists: 2017 Naomi Nakao Gender-Based Research Award: 1081. *Official journal of the American College of Gastroenterology| ACG, 112*, S593-S594.

Butler, R. J., & Johnson, W. G. (2016). Rating the digital help: electronic medical records, software providers, and physicians. *Int J Health Econ Manag, 16*(3), 269-283. doi:10.1007/s10754-016-9190-8

Cerny, B. A., & Kaiser, H. F. (1977). A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate behavioral research, 12*(1), 43-47.

Contratto, E., Romp, K., Estrada, C. A., Agne, A., & Willett, L. L. (2017). Physician Order Entry Clerical Support Improves Physician Satisfaction and Productivity. *South Med J, 110*(5), 363-368. doi:10.14423/smj.0000000000000645

Copley, L. A., Sharps, C. H., Gerardi, J. A., Gupta, S. K., Vanderhaave, K. L., Lovejoy, J. F., . . . Jo, C. (2019). Electronic medical record use and satisfaction among pediatric orthopaedic surgeons. *Journal of Pediatric Orthopaedics, 39*(9), e722-e728. Retrieved from https://www.ingentaconnect.com/content/wk/bpo/2019/00000039/00000009/art00022;jsessionid=3qnx11g0gy1s.x-ic-live-01

Dewa, C. S., Jacobs, P., Thanh, N. X., & Loong, D. (2014). An estimate of the cost of burnout on early retirement and reduction in clinical hours of practicing physicians in Canada. *BMC Health Services Research, 14*(1), 1-9.

Dewa, C. S., Loong, D., Bonato, S., Thanh, N. X., & Jacobs, P. (2014). How does burnout affect physician productivity? A systematic literature review. *BMC Health Serv Res, 14,* 325. doi:10.1186/1472-6963-14-325

Dolan, E. D., Mohr, D., Lempa, M., Joos, S., Fihn, S. D., Nelson, K. M., & Helfrich, C. D. (2015). Using a single item to measure burnout in primary care staff: a psychometric evaluation. *Journal of general internal medicine, 30*(5), 582-587. doi:10.1007/s11606-014-3112-6

Domaney, N. M., Torous, J., & Greenberg, W. E. (2018). Exploring the Association Between Electronic Health Record Use and Burnout Among Psychiatry Residents and Faculty: a Pilot Survey Study. *Academic psychiatry : the journal of the American Association of Directors of Psychiatric Residency Training and the Association for Academic Psychiatry, 42*(5), 648-652. doi:10.1007/s40596-018-0939-x

Dutta, B., & Hwang, H.-G. (2020). The adoption of electronic medical record by physicians: A PRISMA-compliant systematic review. *Medicine, 99*(8), e19290-e19290. doi:10.1097/MD.0000000000019290

Dymek, C., Kim, B., Melton, G. B., Payne, T. H., Singh, H., & Hsiao, C. J. (2021). Building the evidence-base to reduce electronic health record-related clinician burden. *J Am Med Inform Assoc, 28*(5), 1057-1061. doi:10.1093/jamia/ocaa238

Dziuban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological bulletin, 81*(6), 358.

Fred, H. L., & Scheid, M. S. (2018). Physician Burnout: Causes, Consequences, and (?) Cures. *Tex Heart Inst J, 45*(4), 198-202. doi:10.14503/thij-18-6842

Freudenberger, H. J. (1974). Staff burn-out. *Journal of social issues, 30*(1), 159-165.

Friedberg, M. W., Chen, P. G., Van Busum, K. R., Aunon, F., Pham, C., Caloyeras, J., . . . Brook, R. H. (2014). Factors affecting physician professional satisfaction and their implications for patient care, health systems, and health policy. *Rand health quarterly, 3*(4).

Friedman, C., Rubin, J., Brown, J., Buntin, M., Corn, M., Etheredge, L., . . . Van Houweling, D. (2014). Toward a science of learning systems: a research agenda for the high-functioning Learning Health System. *Journal of the American Medical Informatics Association, 22*(1), 43-50. doi:10.1136/amiajnl-2014-002977

Gamer, M., J. Lemon, I. Fellows, & P. Singh. (2019). *irr: Various Coefficients of Interrater Reliability and Agreement.* R package version 0.84.1. https://CRAN.R-project.org/package=irr

Gardner, R. L., Cooper, E., Haskell, J., Harris, D. A., Poplau, S., Kroth, P. J., & Linzer, M. (2019). Physician stress and burnout: the impact of health information technology. *J Am Med Inform Assoc, 26*(2), 106-114. doi:10.1093/jamia/ocy145

Gil-Monte, P. R., Rojas, S. U., & Ocaña, J. I. S. (2009). Validez factorial del «Cuestionario para la Evaluación del Síndrome de Quemarse por el Trabajo»(CESQT) en una muestra de maestros mexicanos. *Salud mental*, *32*(3), 205-214.

Gilleland, M., Komis, K., Chawla, S., Fernandez, S., Fishman, M., & Adams, M. (2014). Resident duty hours in the outpatient electronic health record era: inaccuracies and implications. *J Grad Med Educ, 6*(1), 151-154. doi:10.4300/jgme-d-13-00061.1

Gregory, M. E., Russo, E., & Singh, H. (2017). Electronic Health Record Alert-Related Workload as a Predictor of Burnout in Primary Care Providers. *Appl Clin Inform*, *8*(3), 686-697. doi:10.4338/aci-2017-01-ra-0003

Grumbach, K., Knox, M., Huang, B., Hammer, H., Kivlahan, C., & Willard-Grace, R. (2019). A Longitudinal Study of Trends in Burnout During Primary Care Transformation. *The Annals of Family Medicine, 17*(Suppl 1), S9-S16. doi:10.1370/afm.2406

Gupta, K., Murray, S. G., Sarkar, U., Mourad, M., & Adler-Milstein, J. (2019). Differences in ambulatory EHR use patterns for male vs. female physicians. *NEJM Catalyst*, *5*(6).

Halbesleben, J. R., & Rathert, C. (2008). Linking physician burnout and patient outcomes: exploring the dyadic relationship between physicians and patients. *Health care management review, 33*(1), 29-39.

Hamade, N., Terry, A., & Malvankar-Mehta, M. (2019). Interventions to improve the use of EMRs in primary health care: a systematic review and meta-analysis. *BMJ health & care informatics, 26*(1), 0-0. doi:10.1136/bmjhci-2019-000023

Han, S., Shanafelt, T. D., Sinsky, C. A., Awad, K. M., Dyrbye, L. N., Fiscus, L. C., . . . Goh, J. (2019). Estimating the Attributable Cost of Physician Burnout in the United States. *Ann Intern Med, 170*(11), 784-790. doi:10.7326/m18-1422

Harris, D. A., Haskell, J., Cooper, E., Crouse, N., & Gardner, R. (2018). Estimating the association between burnout and electronic health record-related stress among advanced practice registered nurses. *Appl Nurs Res, 43*, 36-41. doi:10.1016/j.apnr.2018.06.014

Hash, J., Bowen, P., Johnson, A., Smith, C., & Steinberg, D. (2005). *An introductory resource guide for implementing the health insurance portability and accountability act (HIPAA) security rule*: US Department of Commerce, Technology Administration, National Institute of ….

Hauer, A., Waukau, H. J., & Welch, P. (2018). Physician Burnout in Wisconsin: An Alarming Trend Affecting Physician Wellness. *Wmj, 117*(5), 194-200.

Hilliard, R. W., Haskell, J., & Gardner, R. L. (2020). Are specific elements of electronic health record use associated with clinician burnout more than others? *J Am Med Inform Assoc*. doi:10.1093/jamia/ocaa092

Hoff, T., Whitcomb, W. F., & Nelson, J. R. (2002). Thriving and surviving in a new medical career: the case of hospitalist physicians. *Journal of Health and Social Behavior, 43*(1), 72.

Horwitz, L. I., Kuznetsova, M., & Jones, S. A. (2019). Creating a Learning Health System through Rapid-Cycle, Randomized Testing. *N Engl J Med*, *381*(12), 1175-1179. doi:10.1056/NEJMsb1900856

InformedHealth.org [Internet]. (2020, June 18). *Depression: What is burnout?* Institute for Quality and Efficiency in Health Care (IQWiG) https://www.ncbi.nlm.nih.gov/books/NBK279286/.

Institute of Medicine. Learning Healthcare Project. Retrieved November 18, 2020, from http://www.learninghealthcareproject.org/section/background/learning-healthcare-system.

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika, 35*(4), 401-415.

Karp, E. L., Freeman, R., Simpson, K. N., & Simpson, A. N. (2019). Changes in Efficiency and Quality of Nursing Electronic Health Record Documentation After Implementation of an Admission Patient History Essential Data Set. *Comput Inform Nurs, 37*(5), 260-265. doi:10.1097/cin.0000000000000516

Khairat, S., Burke, G., Archambault, H., Schwartz, T., Larson, J., & Ratwani, R. M. (2018). Perceived Burden of EHRs on Physicians at Different Stages of Their Career. *Appl Clin Inform, 9*(2), 336-347. doi:10.1055/s-0038-1648222

KLAS. Arch Collaborative. Retrieved November 18, 2020, from https://klasresearch.com/arch-collaborative.

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*, *15*(2), 155-163. doi:10.1016/j.jcm.2016.02.012

Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work & Stress, 19*(3), 192-207.

Kroth, P. J., Morioka-Douglas, N., Veres, S., Babbott, S., Poplau, S., Qeadan, F., . . . Linzer, M. (2019). Association of Electronic Health Record Design and Use Factors With Clinician Stress and Burnout. *JAMA Netw Open, 2*(8), e199609. doi:10.1001/jamanetworkopen.2019.9609

Malakh-Pines, A., Aronson, E., & Kafry, D. (1981). *Burnout: From tedium to personal growth*: New york: free press.

Marckini, D. N., Samuel, B. P., Parker, J. L., & Cook, S. C. (2019). Electronic health record associated stress: A survey study of adult congenital heart disease specialists. *Congenit Heart Dis, 14*(3), 356-361. doi:10.1111/chd.12745

Maslach, C., Jackson, S. E., Leiter, M. P., Schaufeli, W. B., & Schwab, R. L. (1986). *Maslach burnout inventory* (Vol. 21): Consulting psychologists press Palo Alto, CA.

Maslach, C., Jackson, S. E., Leiter, M. P., Schaufeli, W. B., & Schwab, R. L. (2016). Maslach Burnout Inventory [Fourth Edition Manual]. In M. G. Kavan & D. Powell (Eds.), *MBI*.

Maslach, C., & Leiter, M. P. (2014). Editorial. *Burnout Research, 1*(1), 1-2. doi:https://doi.org/10.1016/j.burn.2014.03.001

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual review of psychology, 52*(1), 397-422.

McGinnis, J. M., Stuckhardt, L., Saunders, R., & Smith, M. (2013). *Best care at lower cost: the path to continuously learning health care in America*: National Academies Press.

Mehta, L. S., Lewis, S. J., Duvernoy, C. S., Rzeszut, A. K., Walsh, M. N., Harrington, R. A., . . . Douglas, P. S. (2019). Burnout and career satisfaction among US cardiologists. *Journal of the American College of Cardiology, 73*(25), 3345-3348.

Melnick, E. R., Dyrbye, L. N., Sinsky, C. A., Trockel, M., West, C. P., Nedelec, L., . . . Shanafelt, T. (2020). The Association Between Perceived Electronic Health Record Usability and Professional Burnout Among US Physicians. *Mayo Clin Proc, 95*(3), 476-487. doi:10.1016/j.mayocp.2019.09.024

Melnick, E. R., Ong, S. Y., Fong, A., Socrates, V., Ratwani, R. M., Nath, B., . . . Sinsky, C. A. (2021). Characterizing physician EHR use with vendor derived data: a feasibility study and cross-sectional analysis. *Journal of the American Medical Informatics Association, 28*(7), 1383-1392. doi:10.1093/jamia/ocab011

Menachemi, N., Powers, T., Au, D. W., & Brooks, R. G. (2010). Predictors of physician satisfaction among electronic health record system users. *Journal for Healthcare Quality, 32*(1), 35-41.

Mennemeyer, S. T., Menachemi, N., Rahurkar, S., & Ford, E. W. (2016). Impact of the HITECH Act on physicians' adoption of electronic health records. *J Am Med Inform Assoc, 23*(2), 375-379. doi:10.1093/jamia/ocv103

Micek, M. A., Arndt, B., Tuan, W.-J., Trowbridge, E., Dean, S. M., Lochner, J., . . . Pandhi, N. (2020). Physician burnout and timing of electronic health record use. *ACI Open, 4*(01), e1-e8.

Moy, A. J., Schwartz, J. M., Chen, R., Sadri, S., Lucas, E., Cato, K. D., & Rossetti, S. C. (2021). Measurement of clinical documentation burden among physicians and nurses using electronic health records: a scoping review. *J Am Med Inform Assoc, 28*(5), 998-1008. doi:10.1093/jamia/ocaa325

Murphy, D. R., Meyer, A. N., Russo, E., Sittig, D. F., Wei, L., & Singh, H. (2016). The burden of inbox notifications in commercial electronic health records. *JAMA Internal Medicine, 176*(4), 559-560.

Murphy, P. (2021, April 21). *Exploratory Factor Analysis*. RPubs by RStudio. https://rpubs.com/pjmurphy/758265

Nguyen, O. T., Jenkins, N. J., Khanna, N., Shah, S., Gartland, A. J., Turner, K., & Merlo, L. J. (2021). A systematic review of contributing factors of and solutions to electronic health record-related impacts on physician well-being. *J Am Med Inform Assoc*, *28*(5), 974-984. doi:10.1093/jamia/ocaa339

O'Connor, B. (2015). Achieving the Quadruple Aim. The Triple Aim is great, but there's something it leaves out. *Health Manag Technol*, *36*(6), 24.

Office of the National Coordinator for Health Information Technology. (2019, September) *Percent of Hospitals, By Type, that Possess Certified Health IT, Health IT Quick-Stat #52.* https://www.healthit.gov/data/quickstats/percent-hospitals-type-possess-certified-health-it

Office of the National Coordinator for Health Information Technology. (2019, January). *Office-based Physician Electronic Health Record Adoption, Health IT Quick-Stat #50.* https://www.healthit.gov/data/quickstats/office-based-physician-electronic-health-record-adoption.

Olson, K., Marchalik, D., Farley, H., Dean, S. M., Lawrence, E. C., Hamidi, M. S., . . . Stewart, M. T. (2019). Organizational strategies to reduce physician burnout and improve professional fulfillment. *Curr Probl Pediatr Adolesc Health Care, 49*(12), 100664. doi:10.1016/j.cppeds.2019.100664

Olson, K., Sinsky, C., Rinne, S. T., Long, T., Vender, R., Mukherjee, S., . . . Linzer, M. (2019). Cross-sectional survey of workplace stressors associated with physician burnout measured by the Mini-Z and the Maslach Burnout Inventory. *Stress Health, 35*(2), 157-175. doi:10.1002/smi.2849

Overhage, J. M., & McCallie, D., Jr. (2020). Physician Time Spent Using the Electronic Health Record During Outpatient Encounters: A Descriptive Study. *Ann Intern Med, 172*(3), 169-174. doi:10.7326/m18-3684

Panagioti, M., Geraghty, K., Johnson, J., Zhou, A., Panagopoulou, E., Chew-Graham, C., . . . Esmail, A. (2018). Association Between Physician Burnout and Patient Safety, Professionalism, and Patient Satisfaction: A Systematic Review and Meta-analysis. *JAMA Internal Medicine, 178*(10), 1317-1331. doi:10.1001/jamainternmed.2018.3713

Patel, R. S., Bachu, R., Adikey, A., Malik, M., & Shah, M. (2018). Factors Related to Physician Burnout and Its Consequences: A Review. *Behavioral sciences (Basel, Switzerland), 8*(11), 98. doi:10.3390/bs8110098

Pinevich, Y., Clark, K. J., Harrison, A. M., Pickering, B. W., & Herasevich, V. (2021). Interaction Time with Electronic Health Records: A Systematic Review. *Appl Clin Inform, 12*(4), 788-799. doi:10.1055/s-0041-1733909

Poon, E. G., Trent Rosenbloom, S., & Zheng, K. (2021). Health information technology and clinician burnout: Current understanding, emerging solutions, and future directions. *J Am Med Inform Assoc, 28*(5), 895-898. doi:10.1093/jamia/ocab058

Pozdnyakova, A., Laiteerapong, N., Volerman, A., Feld, L. D., Wan, W., Burnet, D. L., & Lee, W. W. (2018). Impact of medical scribes on physician and patient satisfaction in primary care. *Journal of general internal medicine*, *33*(7), 1109-1115.

Privitera, M. R., Atallah, F., Dowling, F., Gomez-DiCesare, C., Hengerer, A., Arnhart, K., . . . Staz, M. (2018). Physicians' electronic health records use at home, job satisfaction, job stress and burnout. *J Hosp Adm*, *7*(4), 52-59.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rassolian, M., Peterson, L. E., Fang, B., Knight, H. C., Peabody, M. R., Baxley, E. G., & Mainous, A. G. (2017). Workplace factors associated with burnout of family physicians. *JAMA Internal Medicine*, *177*(7), 1036-1038.

Ratanawongsa, N., Barton, J. L., Lyles, C. R., Wu, M., Yelin, E. H., Martinez, D., & Schillinger, D. (2016). Association Between Clinician Computer Use and Communication With Patients in Safety-Net Clinics. *JAMA Intern Med*, *176*(1), 125-128. doi:10.1001/jamainternmed.2015.6186

Revelle, W. (2020) *psych: Procedures for Personality and Psychological Research.* Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 2.1.3,.U.S. Department of Health & Human Services. (2020). *Strategy on Reducing the Regulatory and Administrative Burden Relating to the Use of Health IT and EHRs.* https://www.healthit.gov/sites/default/files/page/2020-02/BurdenReport_0.pdf.

Rinne, S. T., Mohr, D. C., Swamy, L., Blok, A. C., Wong, E. S., & Charns, M. P. (2020). National Burnout Trends Among Physicians Working in the Department of Veterans Affairs. *Journal of general internal medicine, 35*(5), 1382-1388. doi:10.1007/s11606-019-05582-7

Robertson, S. L., Robinson, M. D., & Reid, A. (2017). Electronic Health Record Effects on Work-Life Balance and Burnout Within the I(3) Population Collaborative. *J Grad Med Educ, 9*(4), 479-484. doi:10.4300/jgme-d-16-00123.1

Rohland, B. M., Kruse, G. R., & Rohrer, J. E. (2004). Validation of a single-item measure of burnout against the Maslach Burnout Inventory among physicians. *Stress and Health, 20*(2), 75-79. doi:10.1002/smi.1002

Rossetti, S. C., Rosenbloom, F. S. T., Detmer, D., Kevin Johnson, M., Cato, K., Cohen, D., . . . Jessica Schwartz, B. (2021) Report from the 25 By 5: Symposium Series to Reduce Documentation Burden on US Clinicians by 75% by 2025.

Rotenstein, L. S., Torre, M., Ramos, M. A., Rosales, R. C., Guille, C., Sen, S., & Mata, D. A. (2018). Prevalence of Burnout Among Physicians: A Systematic Review. *Jama, 320*(11), 1131-1150. doi:10.1001/jama.2018.12777

Ruitenburg, M. M., Frings-Dresen, M. H. W., & Sluiter, J. K. (2012). The prevalence of common mental disorders among hospital physicians and their association with self-reported work ability: a cross-sectional study. *BMC Health Services Research, 12*(1), 292-298. doi:10.1186/1472-6963-12-292

Rule, A., Chiang, M. F., & Hribar, M. R. (2020). Using electronic health record audit logs to study clinical activity: a systematic review of aims, measures, and methods. *J Am Med Inform Assoc, 27*(3), 480-490. doi:10.1093/jamia/ocz196

Saag, H. S., Shah, K., Jones, S. A., Testa, P. A., & Horwitz, L. I. (2019). Pajama time: working after work in the electronic health record. *Journal of general internal medicine, 34*(9), 1695-1696.

Sauro, J., & Dumas, J. S. (2009). *Comparison of three one-question, post-task usability questionnaires.* Paper presented at the Proceedings of the SIGCHI conference on human factors in computing systems.

Schaufeli, W. B., Leiter, M. P., & Maslach, C. (2010). Burnout: 35 Years of research and practice. *IEEE Engineering Management Review, Engineering Management Review, IEEE, IEEE Eng. Manag. Rev., 38*(4), 4-18. doi:10.1109/EMR.2010.5645750

Scholl, M. (2008). *An introductory resource guide for implementing the Health Insurance Portability and Accountability Act (HIPAA) Security Rule. [electronic resource]* (Rev. 1. ed.): U.S. Dept. of Commerce, National Institute of Standards and Technology.

Semanik, M. G., Kleinschmidt, P. C., Wright, A., Willett, D. L., Dean, S. M., Saleh, S. N., . . . Buchanan, J. R. (2021). Impact of a problem-oriented view on clinical data retrieval. *J Am Med Inform Assoc, 28*(5), 899-906. doi:10.1093/jamia/ocaa332

Shanafelt, T. D., Boone, S., Tan, L., Dyrbye, L. N., Sotile, W., Satele, D., . . . Oreskovich, M. R. (2012). Burnout and Satisfaction With Work-Life Balance Among US Physicians Relative to the General US Population. *Archives of Internal Medicine, 172*(18), 1377-1385. doi:10.1001/archinternmed.2012.3199

Shanafelt, T. D., Dyrbye, L. N., Sinsky, C., Hasan, O., Satele, D., Sloan, J., & West, C. P. (2016). Relationship Between Clerical Burden and Characteristics of the Electronic Environment With Physician Burnout and Professional Satisfaction. *Mayo Clin Proc, 91*(7), 836-848. doi:10.1016/j.mayocp.2016.05.007

Shanafelt, T. D., Hasan, O., Dyrbye, L. N., Sinsky, C., Satele, D., Sloan, J., & West, C. P. (2015). Changes in Burnout and Satisfaction With Work-Life Balance in Physicians and the General US Working Population Between 2011 and 2014. *Mayo Clin Proc, 90*(12), 1600-1613. doi:10.1016/j.mayocp.2015.08.023

Shimotsu, S., Poplau, S., & Linzer, M. (2015). Validation of a brief clinician survey to reduce clinician burnout. *Journal of general internal medicine, 30*(2 suppl), S79-S80.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin, 86*(2), 420.

Sinsky, C., Colligan, L., Li, L., Prgomet, M., Reynolds, S., Goeders, L., . . . Blike, G. (2016). Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties. *Ann Intern Med, 165*(11), 753-760. doi:10.7326/m16-0961

Sinsky, C. A., Rule, A., Cohen, G., Arndt, B. G., Shanafelt, T. D., Sharp, C. D., . . . Hribar, M. (2020). Metrics for assessing physician activity using electronic health record log data. *J Am Med Inform Assoc, 27*(4), 639-643. doi:10.1093/jamia/ocz223

Soler, J. K., Yaman, H., Esteva, M., Dobbs, F., Asenova, R. S., Katić, M., . . . Ungan, M. (2008). Burnout in European family doctors: the EGPRN study. *Family Practice, 25*(4), 245-265. doi:10.1093/fampra/cmn038

Somerson, J. S., Patton, A., Ahmed, A. A., Ramey, S., & Holliday, E. B. (2020). Burnout Among United States Orthopaedic Surgery Residents. *J Surg Educ, 77*(4), 961-968. doi:10.1016/j.jsurg.2020.02.019

Tai-Seale, M., Dillon, E. C., Yang, Y., Nordgren, R., Steinberg, R. L., Nauenberg, T., . . . Frosch, D. L. (2019). Physicians' Well-Being Linked To In-Basket Messages Generated By Algorithms In Electronic Health Records. *Health Aff (Millwood), 38*(7), 1073-1078. doi:10.1377/hlthaff.2018.05509

Tai-Seale, M., Olson, C. W., Li, J., Chan, A. S., Morikawa, C., Durbin, M., . . . Luft, H. S. (2017). Electronic Health Record Logs Indicate That Physicians Split Time Evenly Between Seeing Patients And Desktop Medicine. *Health Aff (Millwood), 36*(4), 655-662. doi:10.1377/hlthaff.2016.0811

Tawfik, D. S., Phibbs, C. S., Sexton, J. B., Kan, P., Sharek, P. J., Nisbet, C. C., . . . Profit, J. (2017). Factors Associated With Provider Burnout in the NICU. *Pediatrics, 139*(5). doi:10.1542/peds.2016-4134

Tedesco, D., & Tullis, T. (2006). A comparison of methods for eliciting post-task subjective ratings in usability testing. *Usability Professionals Association (UPA), 2006,* 1-9.

Trockel, M., Bohman, B., Lesure, E., Hamidi, M. S., Welle, D., Roberts, L., & Shanafelt, T. (2018). A Brief Instrument to Assess Both Burnout and Professional Fulfillment in Physicians: Reliability and Validity, Including Correlation with Self-Reported Medical Errors, in a Sample of Resident and Practicing Physicians. *Academic psychiatry : the journal of the American Association of Directors of Psychiatric Residency Training and the Association for Academic Psychiatry, 42*(1), 11-24. doi:10.1007/s40596-017-0849-3

Tsou, A. Y., Lehmann, C. U., Michel, J., Solomon, R., Possanza, L., & Gandhi, T. (2017). Safe Practices for Copy and Paste in the EHR. Systematic Review, Recommendations, and Novel Model for Health IT Collaboration. *Applied clinical informatics*, *8*(1), 12-34. doi:10.4338/ACI-2016-09-R-0150

Wallace, J. E., Lemaire, J. B., & Ghali, W. A. (2009). Physician wellness: a missing quality indicator. *Lancet, 374*(9702), 1714-1721. doi:10.1016/s0140-6736(09)61424-0

Ward, Z. (2019). *Physician Stress: Is the Electronic Health Record to Blame?* Western Kentucky University,

Weintraub, A., Geithner, E., Stroustrup, A., & Waldman, E. (2016). Compassion fatigue, burnout and compassion satisfaction in neonatologists in the US. *Journal of Perinatology, 36*(11), 1021-1026.

West, C. P., Dyrbye, L. N., Erwin, P. J., & Shanafelt, T. D. (2016). Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. *Lancet, 388*(10057), 2272-2281. doi:10.1016/s0140-6736(16)31279-x

West, C. P., Dyrbye, L. N., Sloan, J. A., & Shanafelt, T. D. (2009). Single item measures of emotional exhaustion and depersonalization are useful for assessing burnout in medical professionals. *Journal of general internal medicine, 24*(12), 1318-1321. doi:10.1007/s11606-009-1129-z

Wilkinson, K., Seo, K., Pierce, R., Tonellato, P., Kim, J. H., & Myers, D. (2021, 9-12 Aug. 2021). *Electronic Medical Record Specialty Group Comparison by Multinomial Logistic Regression.* Paper presented at the 2021 IEEE 9th International Conference on Healthcare Informatics (ICHI).

World Health Organization. (2019, May 28). *Burn-out an "occupational phenomenon": International Classification of Diseases"*. https://www.who.int/news/item/28-05-2019-burn-out-an-occupational-phenomenon-international-classification-of-diseases.

Wylie, M. C., Baier, R. R., & Gardner, R. L. (2014). Perceptions of electronic health record implementation: a statewide survey of physicians in Rhode Island. *The American journal of medicine, 127*(10), 1010. e1021-1010. e1027.

Yan, Q., Jiang, Z., Harbin, Z., Tolbert, P. H., & Davies, M. G. (2021). Exploring the relationship between electronic health records and provider burnout: A systematic review. *J Am Med Inform Assoc, 28*(5), 1009-1021. doi:10.1093/jamia/ocab009

Zhang, W., Gunter, C. A., Liebovitz, D., Tian, J., & Malin, B. (2011). Role prediction using Electronic Medical Record system audits. *AMIA Annu Symp Proc, 2011*, 858-867.

VITA

Katie Wilkinson is the Sr. Director responsible for data management, application development and research at the Tiger Institute for Health Innovation, a public-private partnership between the University of Missouri and Cerner Corporation, a healthcare information technology (HIT) vendor based in Kansas City, MO. Her research focuses on improving HIT and use of data for and analytics for research and clinical care.

Katie was awarded the Nasa Langley Undergraduate Student Research Project (USRP) scholarship from 1994-1998 and received her bachelor's degrees in political science and mathematics with an emphasis in computer science in 1998 from the University of Northern Colorado. Katie was awarded the Madison and Lila Self graduate fellow at the University of Kansas and completed her master's degree in political science. She joined the MU Institute for Data Science and Informatics in 2017 and completed her PhD research on an Evaluation of an Electronic Medical Record (EMR) Based System to Characterize and Correlate Physician Burnout and EMR Use.