Examining Factors Associated with Learning and Performance in Primary Care Graduate Medical Education Organizations

by

Jung Gook Kim

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Health Policy

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Hector P. Rodriguez, Chair
Professor Stephen M. Shortell
Professor Bruce Fuller
Professor Diane R. Rittenhouse

Summer 2019

ABSTRACT

Examining Factors Associated with Learning and Performance in Primary Care Graduate Medical Education Organizations

by

Jung Gook Kim

Doctor of Philosophy in Health Policy

University of California, Berkeley

Professor Hector P. Rodriguez, Chair

Despite calls to improve Graduate Medical Education (GME), little is known about the organizational factors influencing training design, resident learning, and assessment. This dissertation examines the organizational behavior factors in primary care GME associated with time spent training in ambulatory care, resident clinical competency learning rates, and quality of care. Linked databases from medical education accreditors and policymakers, population health sources, federal cost reports, and an integrated health system were analyzed to investigate the extent to which primary care GME's competing internal and external organizational factors influence the professional training environment and performance of primary care residents. Key findings include: 1. Experience in ambulatory care for residents varies among their ACGME-accredited programs, with more time in ambulatory care settings most strongly associated with additional faculty, receipt of federal Teaching Health Center GME funding, and accreditation warnings; 2. Improved resident learning rates in the Accreditation Council for GME (ACGME) Milestones for family medicine and internal medicine programs were more associated with external factors than internal factors. Patient care, practice-based learning and improvement, and systems-based practice learning rates were dependent on the program's geographic setting, organizational structural characteristics, and the type of resident learning experiences; and 3. Healthcare Effectiveness Data and Information Set (HEDIS) measure reliability in ambulatory care for residents varies among ACGME-accredited primary care residency programs with potential opportunities to utilize publicly reported quality data for GME programs. Overall, these empirical studies help clarify the organizational and associated environmental factors influencing training design, resident learning, and performance in order to assist policymakers in understanding the fragmented GME learning environment and move GME toward improved accountability. As trainee experiences may have a downstream impact on patient care, the systematic study on primary care GME organizations helps improve the design of the resident learning environment and training the future primary care workforce, especially in ambulatory care, the most common delivery setting for primary care health services today.

DEDICATION

This dissertation is in honor of all anchor parents,
in recognition of all the heroic acrobats required of plate spinning…

Be kind, grateful, and work hard. Repeat.

```
___/   /   /___/   /   /___/   15.1   Copyright 1985-2017 StataCorp LLC
  Statistics/Data Analysis            StataCorp
                                      4905 Lakeway Drive
     MP - Parallel Edition            College Station, Texas 77845 USA
                                      800-STATA-PC        http://www.stata.com
                                      979-696-4600        stata@stata.com
                                      979-696-4601 (fax)

Single-user 4-core Stata perpetual license:
       Serial number:  50150
         Licensed to:  JG Kim
                       UC Berkeley
```

Command

```
*20/May/2019
clear all
log using "healthpolicyphd_program_pearls.log",replace

use "healthpolicy_programtenure.dta"

sort time priorities choices

forvalues i = 1/365 {
gen hardwork`i'=1 if motivation !=0
replace hardwork`i'=0 if motivation==0
gen kindness`i'=1 if fulfillment !=0
replace kindness`i'=0 if fulfillment==0
gen gratitude`i'=1 if entitlement !=0
replace gratitude`i'=0 if entitlement==0
}

list

forvalues i = 1/365 {
gen success`i'=1 if hardwork`i'==1 & kindness`i'==1 & gratitude`i'==1
}

list

save "healthpolicy_PhD.dta"

log close
exit, clear
```

ACKNOWLEDGEMENTS

My doctorate journey was undeniably a village effort:

Foremost, to Laurel. My touchstone and sage. You light each stone for every step I take.

To my nuclear unit: My husband, Mr. Farnsworth who makes the mundane everyday invisible so I'm able to shine. I take solace in your unstated brilliance and aversion for unnecessary ambition. To our children: Darjeeling, your resilience is a testament to health and embodies the gifts of long-term choices. Heathrow, that unconditional love is never out of reach. And to our late Barclay, that fulfillment in life is critical when any day may be your last. To 마리아 : for ushering the family into the republic of California and making the impossible possible. To my parents: Mom, your deference, steadfastness, and nourishment are foundational tools that drive my every decision. To my late father, who always reminded to live in gusto and never leave a stone unturned.

To my Group Health Family: Carl and Cindy: my yetis. Your wisdom is awe inspiring and grounding. Andrea and Kasia: Thank you for all the pulse checks. Lisa & Jason: a team like us can never be assumed – Destiny's Child will always live on. To my Kaiser Permanente Family: Lindsay, Nardine, Jen, Dan, Bruce, Michael, Danny, Marc, Nancy, Chris, Paul, Abbas, and Carla: Thank you for cheerleading from the sidelines and always providing a learning environment I could thrive in - I look forward to our new adventures.

To my co-ladies in the Accountable Ladies Organization: Your fellowship has generated a lifetime of memories, and your humaneness nourishes me continuously. Aryn: your ability to redirect emotional polarity from negative to positive has energized me over these three years, and especially during the most disparaging of times. Leeann: your courage and how you achieve success by modeling through action and choice words makes me feel honored to walk beside you, as we make sense of this other side of the glass ceiling.

To my Health Policy classmates and cohort: co-learning with you has constantly reminded to value my own growth and find gratitude to what I have achieved. David especially, you are a reminder that humbled brilliance without being hypercompetitive and coupled with gratitude is always the path worth taking.

To CHOIR: My "big sisters" Kathy, Vanessa, Jess. Thank you for always being generous with your time and sharing unconditionally your own paths of development. To Salma, Zosha, Dora, and Patty. Your grace and unwavering support were never unnoticed.

To my GSE and Epi friends: To Bruce's Writing Group – I will always value your insights that will shape education policy and help improve learning for our future generations. To Erika: You are a sunflower in this jungle. I only regret meeting you half-way into my time at Berkeley.

To the School of Public Health: The brilliant faculty that taught me pearls for a lifetime: Will, Joan, Ann, and Tim. To Marques and Kristina: thank you for being our ushers to the massive complex institution of UC Berkeley. To my Joint Medical Program family: Amin, Susan, John, and Gustavo. Thank you for giving me a MedEd home. And to Deborah: Thank you for allowing me to loiter and ensuring OOMPH was never far away.

To the AAMC, ACGME, and Kaiser Permanente for supporting my studies on GME, aligning medical education to health systems, and helping us educators understand how our future workforce can grow and prosper in the unchartered waters ahead.

Last but not least, to my dissertation advisors and mentors:

- To my Chair, Hector: Thank you for taking in this rough diamond and brillianteering my professional prospects. Your transparency with your own journey is forever appreciated.
- To Steve: Thank you for embodying and continuously advocating for Public Health. Your uncanny ability to foresee the future is unmatched and I'm eternally grateful to have learned under your mastery of health services. Your unwavering support has been a lifeline.
- To Bruce: Thank you for illuminating that scholars and their research can be both practical and rigorous. Inviting me into your education home has provided a lifetime of fruits. And thank you for always reminding me that rest and refueling are key to scholarship.
- To Eric: Thank you for being my GME attaché and entrusting me with your decades-long hard earned efforts that align education with health services. This is only the beginning.
- To Diane: Thank you for your maternal spirit coupled with professional grace. You are a beacon of hope by illuminating that work-life balance is indeed achievable in today's society. Your bravery and poise forge my next professional steps. I will always ask myself, "What would Diane do?"

Thank you all. I am forever grateful. Fiat Lux.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

List of Tables

List of Figures

LIST OF ABBREVIATIONS

AAMC	Association of American Medical Colleges
ABMS	American Board of Medical Specialties
ACA	Affordable Care Act
ACGME	Accreditation Council for Graduation Medical Education
AHRF	Area Health Resource File
AIC	Akaike Information Criterion
AMA	American Medical Association
ANOVA	Analysis-of-variance
BIC	Bayesian Information Criterion
CBME	Competency-based Medical Education
CCC	Clinical Competency Committee
CMS	Centers for Medicare and Medicaid Services
DIO	Designated Institutional Officer
FIPS	Federal Information Processing Standards
FM	Family Medicine
FPL	Federal Poverty Level
GME	Graduate Medical Education
HbA1c	Hemoglobin A1c
HEDIS	Healthcare Effectiveness Data and Information Set
HPSA	Health Professions Shortage Area
HRSA	Health Resources and Services Administration
ICC	Intraclass-correlation
IM	Internal Medicine
KP	Kaiser Permanente
KPSCAL	Kaiser Permanente Southern California
LDL-C	Low-density lipoprotein cholesterol
NAM	National Academy of Medicine
NBME	National Board of Medical Examiners
NCQA	National Committee for Quality Assurance
PGY	Postgraduate Year
PGY1	Postgraduate Year 1st Year Resident
RRC	Residency Review Committees
SI	Sponsoring Institution
THCGME	Teaching Health Center Graduate Medical Education
VIF	Variance Inflation Factors

CHAPTER 1: INTRODUCTION

Graduate Medical Education current state

Graduate Medical Education (GME) is a $16 billion United States (U.S.) workforce training program that produces the nation's practicing physicians.[1,2] Funded primarily by Centers for Medicare and Medicaid Services (CMS), GME training lacks centralized planning and accountability. Consequently, policymakers are highly motivated to reform GME in part of the Patient Protection and Affordable Care Act's (ACA) aim to reform primary care services. Yet even in 2018, the National Academy of Medicine (NAM) reported that GME inadequately delivers its goals to meet the needs of 21st century primary care.[3,4] Their findings include: GME organizational design is fragmented and the field does not function as a national system; the GME system is comprised of independent actors without accountability and make autonomous decisions despite receiving federal funding; and GME training outcomes lacks evidence to systematically train clinically competent physicians, which directly impact patient outcomes.

Competing organizational factors in primary care GME

The genesis of GME organizations in 1928 encompassed dual aims: 1. Residents' service to sponsoring institutions who provide a health care delivery system and patients, and 2. Residents' learning in tandem along with their service requirement to become competent and unsupervised physicians.[5] Over time, GME's dual aims now compete, impacted by the changing delivery system settings from hospitals to ambulatory care, resulting in a fragmented learning environment.[5–10] The majority of sponsoring institutions remain funded by CMS in hospital-based settings and are incentivized to preserve these settings.[1,11] Despite the learning environment needing to mirror where physicians will most likely practice, residents only spent 36% of their entire training in their own medical specialty because of this misalignment.[12–14] Consequently, for the two largest medical specialties in primary care, residents in family medicine and internal medicine describe their ambulatory care training as inadequate and stressful due to limited exposure in the outpatient setting, lacking patient continuity, and cognitively distracting due to competing demands with hospital-based training.[15–17]

Further complicating GME training are recent policy shifts that structure the GME learning environment.[5] GME's accrediting body, the Accreditation Council for GME (ACGME) now requires competency-based medical education (CBME) to teach and assess whether residents are adequately prepared for unsupervised practice.[13,18] To stay accredited, GME programs must provide a formal learning environment to systematically assess for physician clinical competency.[19] Thus decisions in structuring the GME learning environment is impacted by a multitude of factors, internal to the organization to meet resident learning needs and a program's sponsoring institutions' incentives, and external with accrediting bodies and other environmental factors including dependence on CMS-funding.

Yet little knowledge is available on these competing GME organizational factors and impact on learning in ambulatory care.[20–23] Thus a key opportunity is presented to study GME's current state by examining the competing organizational and environmental factors and their association with decision-making to the design and assessment in the resident learning environment.[24]

CHAPTER 2: CONCEPTUAL MODEL & SPECIFIC AIMS

How do theories of organization behavior clarify factors in primary care GME organizations that impact program educational choices and competence for resident physicians?

By incorporating organizational behavior theories on institutional logics, organizational design, resource dependency, and organizational learning, this dissertation investigates the extent to which primary care GME's competing organizational factors influence the professional training environment and technical competence of primary care physicians.

Institutional logics

To understand the competing tensions with GME organizations, institutional logics, or the "assumptions, values, and beliefs, by which … organizations … organize time and space" provide insight to the decision making impacted by competing organizational factors.[25] Multiple logics, or the plurality of symbols and materials in organizations managed simultaneously are common in fragmented fields and compete for attention and time, thus impacting decision making.[25–27] A previous study by Dunn and Jones in medical schools found that competing factors create a "dynamic tension" and impacts the learning of future physicians.[28] Hence, logics materialized as internal and external organizational attributes in GME could illuminate the factors influencing decision-making across GME learning environments.

Organizational Design & Resource Dependency

Evidence for organizational design argues that organizational structures are not always rational and reflect underlying political dynamics and complexity from external to the organization.[29–31] Additionally, how organizations respond to their external environment may reflect how they are structured due to their dependence on resources.[32–34] Consequently, external factors via environmental resources influenced by CMS-funding may be antecedent macro organizational factors in the form of how many Medicare enrollees, uninsured individuals, and amount of available primary care services are in geographic proximity to which GME programs and their sponsoring institutions are set in. Thus, it is plausible to hypothesize that ambulatory care training design is dependent on the environmental context in which the GME organization operates under. This could impact decision making for internal organizational design factors including differentiation of training experiences, size of programs, and life cycle or survival of GME programs and sponsoring institutions. Thus by studying the structural fragmentation observed in the GME organizational design landscape, this could help better understand the competing demands within GME due to their external and internal organizational factors.

Organizational Learning

How would competing organizational factors impact resident learning within their settings? The organizational learning literature argues that organizations cultivate knowledge, behaviors and attitudes to target new ways of behaving and performing within their technical core.[35–37] Consequently, organizations faced with competing factors both internally and externally would present differing strategies and routines, impacting the cumulative experiences of those within the organization. This could also lead to competency traps, whereas an organization could produce maladaptive routines.[35] For GME organizations, an anticipated outcome would be a varying learning experience for residents, based on organizational factors associated with the

ACGME-core competencies that measure residents' competence in the micro environment. Thus over time, how this impacts residents' learning and their achievement of clinical competency as defined by ACGME match calls to understanding how learning in GME programs aligns with the health systems they operate under.[35,38–40]

Figure 2.1 depicts a conceptual model that integrates these organizational behavior theories with current empirical literature. By utilizing linked databases from the ACGME, the Association of American Medical Colleges (AAMC), the Area Health Resource File (AHRF), the Centers for Medicare & Medicaid Services (CMS) cost reports, the Robert Wood Johnson Foundation/University of Wisconsin Population Health Institute County Health Rankings, and the Kaiser Permanente (KP) Quality of Care indicators, this dissertation's objective will study the following three aims for residents training in family medicine and internal medicine accredited GME programs:

Aim 1: Factors Associated with Training Time in Ambulatory Care for First Year Residents in Family Medicine and Internal Medicine

- What internal organizational and external environmental factors are associated with training time in ambulatory care for first year residents in family medicine and internal medicine?

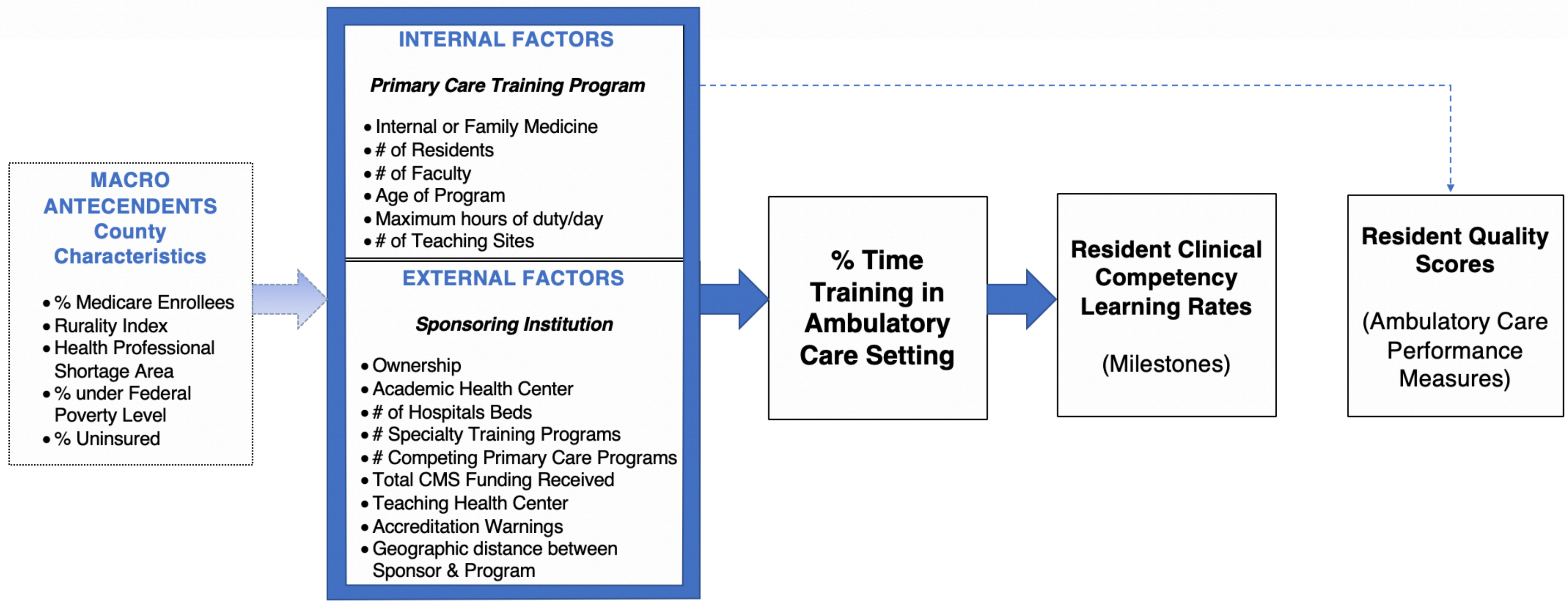
Aim 2: Organizational Factors Associated with Improved Core Competencies among Family Medicine and Internal Medicine Residents

- What predictors for internal organizational and external environmental factors impact the cumulative learning rates for primary care residents training in ACGME-accredited Family Medicine and Internal Medicine programs?

Aim 3: The Reliability of Resident Performance training in Family Medicine and Internal Medicine programs

- What is the reliability of ambulatory care quality scores for family medicine and internal medicine residents if we examine their performance on publicly reported data during their formal training periods?

Figure 2.1: Overall Conceptual Model



CHAPTER 3: AIM 1

Factors Associated with Training Time in Ambulatory Care for First Year Residents in Family Medicine and Internal Medicine

Background

In the United States (U.S.), Graduate Medical Education (GME), or residency training, determines the composition and competencies of the physician workforce.[3,4] GME is highly subsidized by the federal government through the Centers for Medicare and Medicaid Services (CMS) programs, which provide the majority of GME funding nationwide.[11] These CMS programs make GME payments to teaching hospitals with the result that residency training mostly occurs in hospital settings and focuses on the care of inpatients. Inpatient-focused training is appropriate for certain medical and surgical specialties, but not for all family medicine and internal medicine residents who will practice mostly in ambulatory care settings upon graduation.[41] However, the Family Medicine and Internal Medicine Residency Review Committees (RRCs) of the Accreditation Council for GME (ACGME) set the program requirements for residency programs and only provide general guidelines for the amount of time spent training in ambulatory care settings.[42,43] The National Academy of Medicine and key stakeholders call for GME funding and accreditation reforms to incentivize the training of family medicine and internal medicine physicians in ambulatory care settings.[3,4] The relatively small federal Teaching Health Center GME program administered through the U.S. Health Resources and Services Administration (HRSA) funds residency training in community-based and rural clinics and holds great promise as an alternative to hospital-based residency training models.[6,44]

Studies on primary care GME describe residents' ambulatory care time as crucial for developing continuity of care with patients.[22,45–48] Increased continuity of care has high value for patients, by enabling early and effective management of health care problems, and reduces unnecessary care and hospitalizations.[49,50] In addition, GME and higher education experts argue early immersion in ambulatory care settings is essential to the development of competency and formation of professional identity for primary care physicians.[47,51–54] In the U.S., decisions about the amount of time residents spend in inpatient versus ambulatory care settings reside with individual residency programs.[55,56] Additionally, previous evidence reports GME organizational factors may impact postgraduate year (PGY) training outcomes yet little is known on whether these factors within the GME organization and environment are associated with the amount of ambulatory care time for first year postgraduate year residents (PGY1s).[1,57]

To shed light on the factors influencing these decisions, this study uses unique national data to describe the primary care GME landscape, and to examine the organizational and environmental factors associated with training time in ambulatory care for PGY1s in family medicine and internal medicine.

Methods

Data

This study is a cross-sectional analysis using integrated national datasets: the 2016-2017 American Medical Association (AMA) and Association of American Medical Colleges (AAMC)'s National GME Census of ACGME-accredited programs, 2016 CMS Cost Reports and the 2015-2016 Area Health Resource File (AHRF).[58,59] GME sponsoring institutions' primary hospitals were linked to CMS Cost Reports, and GME sponsoring institutions' Federal Information Processing Standards (FIPS) county codes were linked to the AHRF. We also identified sponsoring institutions who had active HRSA Teaching Health Center GME funding in 2016 via the American Association of Teaching Health Centers.[60] The GME Census survey is completed annually by all U.S. GME programs, with a 95.7% response rate.[61]

Study Population

We selected the two largest primary care medical specialties, family medicine and internal medicine, for our study sample. We recognized internal medicine programs include an open-ended career pathway to hospital-based training, but incorporated all internal medicine programs due to our outcome variable examining PGY1s.[62] 1,009 accredited residency programs under 612 GME sponsoring institutions identified by the AAMC were incorporated into the analysis. We excluded programs that were not active during the reporting year period. We also excluded programs that were military-funded, and programs located in Puerto Rico, due to our analysis that examined CMS funding and geographic factors, respectively. We also removed programs that did not respond to the outcome variable in the GME Census survey. 895 accredited primary care residency programs nested in 550 sponsoring institutions were included in the study, comprising 88.7% of all ACGME-accredited programs in family medicine and internal medicine nationwide.

Main Outcome Variable

The outcome variable was percent of residents' time spent in the ambulatory care setting during the first year of training. The GME Census survey queries residency programs about the amount of "Training at hospital outpatient clinics" or "Training during first year in ambulatory non-hospital community-based settings, e.g., physician offices, community clinics" for PGY1's.[63] Internal medicine residency programs typically have hospital-based outpatient clinics, whereas family medicine programs typically have ambulatory care clinics that are geographically distinct from the hospital. Therefore for our outcome variable, we relied on the highest percent of ambulatory care time a program reported based on the respondents' intent to report the highest amount of ambulatory care training time.[42,43]

Internal Organizational Characteristics

Internal organizational characteristics of the residency programs included: age; specialty (family medicine vs. internal medicine); number of faculty and residents; reported maximum duty hours/day; and the total number of teaching sites.

External (Environmental) Factors

Environmental factors external to the residency program included characteristics of the sponsoring institution: if sponsoring institutions' ownership type was private/non-profit, set within an academic health center, number of hospital beds, number of (competing) primary care and non-primary care residency programs; total amount of CMS Medicare GME funding received in 2016; and receipt of federal Teaching Health Center funding. In addition, we included whether or not the program had received any accreditation warnings from the ACGME. Finally, we included the distance in miles between the sponsoring institution and the residency program's primary address listed in the GME Census. Distance was calculated using the addresses and geocoordinates for sponsoring institutions and their accredited programs' primary teaching site, using the *georoute* command in STATA.[64] Subsequently, this continuous variable for distance of travel was transformed into a categorical variable due to a large positive skew distribution, and divided into values by 50%ile, 75%ile, 90%ile, and the 95th %ile of its distribution.

Control Variables
Control variables adjusted for sponsoring institution's county characteristics including: whether the entire county was designated as a health professions shortage area (HPSA); the proportion of the county population living below the Federal Poverty Level (FPL); the county's level of rurality using the AHRF's Urban-Rural Continuum Codes; the proportion of the county population that is enrolled in Medicare; and the proportion of the county population that is uninsured.

Statistical Analyses
We compared internal characteristics and external factors for family medicine and internal medicine residency programs using chi-squared for categorical independent variables and t-tests for continuous independent variables. To examine our hypothesis, we used hierarchical linear regression models to estimate the relationship of internal characteristics and external factors with percent time in ambulatory care settings, controlling for county variables. Random sponsor effects were used to account for residency program clustering within sponsoring institutions.[65] We fitted three analytic models (internal factors, external, both internal and external) and tested our independent variables for multicollinearity using variance inflation factors (VIF), which reported a mean VIF of 1.7.[66] We also calculated the intraclass-correlation (ICC) as a post-test to estimate the percent of variance attributed at the program and sponsoring institution's levels.[65] We conducted sensitivity analyses to assess the robustness of our results without clustering at the sponsoring institution through multivariate linear regression and to test only control variables using hierarchical linear models. Results with $p<.05$ were considered statistically significant. Analyses were performed using STATA 15.1 and ArcGIS Pro.[67,68] This study received a human subjects exemption from the University of California, Berkeley Institutional Review Board (IRB).

Results
Residency Program Characteristics
Table 3.1 compares family medicine and internal medicine residency programs according to the percent time in ambulatory care, internal organizational characteristics, external factors, and control variables. The overall mean percent time for first year residents in the ambulatory care setting was 25.4% (SD=12.8), with no significant difference between family medicine and internal medicine.

Compared to internal medicine residency programs, family medicine programs were newer, smaller (in terms of both the number of residents and the number of faculty), and more likely to have received an accreditation warning from the ACGME. Family medicine residency programs were more commonly located in the U.S. Western Census region and in non-metropolitan and rural areas; their sponsoring institution was more likely to have federal Teaching Health Center GME funding and less likely to be an academic health center. Internal medicine programs' sponsoring institutions were more likely to have non-primary care specialty residency programs, receive more CMS Medicare GME funding, be located in the U.S. Northeastern Census region, and be located in a county with a greater proportion of Medicare enrollees.

Overall, the mean distance between the residency program's primary teaching site and their sponsoring institution was 19.7 (SD=3.2) miles. There was no significant difference between family medicine and internal medicine. (Table 3.1)

Figure 3.1 illustrates the national landscape of U.S. Family Medicine and Internal Medicine accredited-GME programs, their sponsoring institutions, and the indication of a program with greater than the reported mean time of 25.4% in the ambulatory care setting.

In adjusted analyses (Table 3.2), first year residents spend a greater percentage of ambulatory care time for every additional faculty member (0.03% (+/-0.01), $p<.001$), if the sponsoring institution received Teaching Health Center funding (6.6% (+/-2.7), $p<.01$) and if the program had accreditation warnings (4.8% (+/-2.5%), $p<.05$). Distance between a program's primary teaching site and sponsoring institution was not statistically significant. For our control variables, less percent of Medicare enrollees in the sponsoring institution's county was associated with more percent time in ambulatory care (see Figure 3.2). While the effect sizes for sponsoring institutions in non-metropolitan and rural counties indicate a greater percent of ambulatory care time, the results were not statistically significant.

External factors including Teaching Health Center GME funding and ACGME accreditation warnings had the strongest association with percent time spent in ambulatory care. The estimated ICC to the variance within sponsoring institutions ranged from 9-14% across models, thus indicating most of the variation in ambulatory care time for first year residents are attributed more at the residency program level than sponsoring institution. Furthermore, sensitivity analyses that modeled linear regression without clustering at the sponsoring institution or hierarchical linear regression using only control variables resulted in consistent results with our main models.

Discussion

Using multiple merged national datasets, we provide a landscape of family medicine and internal medicine residency programs. We identified key factors associated with ambulatory care training time in GME programs and can be used to better identify accountable outcomes.[69]

With regard to our primary outcome, we found that first-year family medicine and internal medicine residents spend about one-quarter of their training time in ambulatory care settings. Currently no benchmark exists for the amount of training time required to establish continuity of care in residency education, despite calls from GME and higher education experts, who argue that early immersion in ambulatory care settings is essential to achieve primary care competency and to mirror primary care physician professional identity.[47,51–54] The calls include GME curricular innovations that prioritize ambulatory care training before all other training needs, as demonstrated by the "Clinic First" movement and other longitudinal continuity clinic immersion models.[22,51,56,70]

Teaching Health Center funding was most strongly associated with increased ambulatory care teaching time for first year family medicine and internal medicine residents – controlling for multiple other internal (organizational) and external factors. These results are timely, given current legislative efforts to extend Teaching Health Center funding beyond October 1, 2019."[71] From our results, it would appear that GME programs in family medicine and internal medicine may feel the burden of the "golden handcuffs" to their sponsoring institutions' dependence on CMS hospital-based Medicare GME funding.[72] Conversely, GME programs federally funded by novel ambulatory care models, like the Teaching Health Centers, appear to be incentivized to structure training toward ambulatory care. Thus, the Teaching Health Center funding effect sizes of 6.6 – 7.2% more time in the ambulatory care setting requires additional investigation given the small size of this subgroup.

More training time in ambulatory care settings was also associated with additional primary care faculty. It may be that programs with more primary care faculty have a greater commitment to ambulatory care training.[6,56] Recruiting dedicated and engaged faculty have been reported as a key factor to optimize ambulatory care training for PGYs.[56] Our finding may also reflect the widely held assumption that teaching first year residents in an ambulatory care setting is more faculty-intensive than teaching in a hospital setting. As GME programs consider the amount of faculty needed, our results suggest that faculty size is a key factor when considering how ambulatory care-based training is designed and funded.

Paradoxically, accreditation warnings from the ACGME were also associated with increased ambulatory care time. One of the most frequent accreditation warnings issued to family medicine programs is the lack of patients cared for in continuity clinics.[73] Thus, a plausible explanation for this finding is that programs are responding to accreditation warnings by increasing training time in the ambulatory care setting. This suggests the importance of setting expectations for accountability, which is a growing argument toward GME outcomes-based deliverables.[3,11,74] An alternative explanation is that programs with more ambulatory care training time receive warnings because of poor training quality. Further study is needed to understand the temporal relationship between accreditation citations and amount of ambulatory care training time, the types of citations received, and how citations play a role in changing training design.[73]

The finding that the proportion of the Medicare enrollees in the sponsoring institution's county is negatively associated with ambulatory care training time. This may imply that residents are primarily in the hospital setting because of their programs sponsored by teaching hospitals, who have more CMS-insured patients or that Medicare patients may be sicker with more complex care needs, resulting in residents to care for these patients. Regardless, this is consistent with the argument that GME funding operates as a "black box" whereas the financial mechanisms to GME funding is unclear and if this incentivizes against ambulatory care training, and requires further study.[75,76]

While other covariates in our models indicate minimal effect sizes, we cannot ignore the potential contribution to assign additional time to patients in the ambulatory care setting. The value of adding 0.5 to 2.0% more time in primary care is high, given the pervasive issue of patients' limited access to primary care, maximizing resident ambulatory care learning time, and mitigate current primary care workforce shortages.[77]

In consideration of our findings, we offer a more nuanced understanding to previous primary care GME studies, GME funding mechanisms, and the institutional decisions that impact the resident educational environment. Primary care GME programs and sponsoring institutions should examine their training design choices for how much time a primary care resident spends in the ambulatory care setting to maximize the utility of existing resident slots for outpatient-based primary care residents to practice in their future settings. If residents are being dissuaded from outpatient-based careers because of feeling inadequate and finding greater comfort in the hospital due to the limited exposure in an ambulatory care setting, an opportunity exists to recommend GME training reforms to mitigate these constraints.[15,56] Our findings suggest the importance of expanding and increasing Teaching Health Center GME funding and more specificity to the RRC requirements.

Limitations

The primary limitation of this study is the self-reported nature of the GME Census survey data, subject to program director or designated institutional officer (DIO) social desirability bias. Additionally, assigning random effects to the sponsoring institution level in our analytic models may not explicitly account for all factors determining residents' time in the ambulatory care setting. While family medicine residents are matched into programs that aim to produce an ambulatory care-based workforce, internal medicine programs are designed with resident preferences to pursue careers in either outpatient or inpatient-based settings and a variety of the American Board of Medical Specialties (ABMS) and fellowships.[61,62] Thus, empirically examining the relationship within primary care GME programs and between their sponsoring institutions should continue to examine the variation in their organizational and environmental factors, and how these factors impact resident education design, clinical competency, related quality of care indicators, and subsequently resident specialty career choices.

Conclusion

Experience in ambulatory care for residents varies among ACGME-accredited primary care residency programs. More time in ambulatory care settings was most strongly associated with receipt of federal Teaching Health Center GME funding and accreditation warnings. Additional faculty was also associated with increased ambulatory care training time. These findings can inform policy discussions on how to increase ambulatory care training time to improve resident learning, competency achievement, and workforce planning in ambulatory care settings.

Table 3.1: Characteristics of Primary Care GME Programs & Sponsoring Institutions

Factor Category	Characteristic	Family Medicine (n=485)	Internal Medicine (n=410)	Overall (n=895)
Internal	**Age of Program (mean, SD)**	31.1 (0.7)	42.7 (1.1) ***	36.5 (0.6)
Internal	**# of Faculty at Program (mean, SD)**	12.7 (0.6)	91.0 (6.5) ***	48.5 (3.3)
Internal	**# Residents at Program (mean, SD)**	21.6 (0.4)	59.1 (2.0) ***	38.9 (1.1)
Internal	Maximum Duty Hours/Week for PGY1 (mean, SD)	16.1 (0.1)	16.5 (0.3)	16.3(0.1)
Internal	**# of teaching sites (mean, SD)**	3.4 (0.1) ***	2.7 (0.1)	3.1 (0.9)
External	**Sponsor is in Academic Health Center (n, %)**	142 (29.3%)	167 (40.7%) ***	309 (34.5%)
External	Sponsor Ownership Type (n, %) Private Non-profit	253 (52.2%)	238 (58.0%)	491 (54.8%)
External	Hospital beds at Sponsor's primary teaching hospital (mean, SD)	543.6 (76.7)	511.2 (16.1)	528 (42.4)
External	**# of Competing Primary Care Program at Same Sponsor (mean, SD)**	1.5 (0.1) **	1.2 (0.1)	1.4 (0.6)
External	**# of Non-Primary Care Programs at Same Sponsor (mean, SD)**	22.5 (1.5)	29.8 (1.8) ***	25.8 (1.2)
External	**Total CMS Medicare GME funding received (mean, SD)**	$6,014,695 ($357,442)	$8,950,660*** ($533,837)	$7,351,933 ($315,121)
External	**Program with accreditation warnings (n, %)**	21 (4.4%) *	6 (1.5%)	27 (3.0%)
External	**SI received Teaching Health Center funding (n, %)**	23 (4.7%) **	5 (1.2%)	28 (3.1%)
External	Distance in miles between Sponsor and Program (mean, SD)	23.1 (3.7)	15.8 (5.5)	19.7 (3.2)
County Characteristics	**% Medicare Enrollees in Sponsor's county (mean, SD)**	16.1(0.0)	16.9(0.0) *	16.6(0.0)
County Characteristics	Sponsor's county is non-metropolitan/rural (n, %)	34 (7.0%)	7 (1.7%)	41 (4.6%)
County Characteristics	Health Prof Shortage Area designation for all sponsor's county (n, %)	188 (38.8%)	145 (35.4%)	333 (37.2%)
County Characteristics	% under FPL in sponsor's county (mean, SD)	15.7 (0.2)	16.2 (0.3)	15.9(0.2)
County Characteristics	Uninsured % in Sponsor's county (mean, SD)	16.2(0.2)	16.0 (0.3)	16.1(0.2)
County Characteristics	Census Region (n, %)			
County Characteristics	**West**	109 (22.5%)***	55 (13.4%)	164 (18.3%)
County Characteristics	Midwest	144 (29.7%)	100 (24.4%)	244 (27.2%)
County Characteristics	South	149 (30.7%)	120 (29.3%)	269 (30.1%)
County Characteristics	**Northeast**	83 (17.1%)	135 (32.9%) ***	218 (24.4%)
Outcome	% of Ambulatory Care time for PGY1 (mean, SD)	25.2 (0.7)	25.7 (0.5)	25.4 (0.4)

**p<.05, **p<.01, ***p<.001*

Table 3.2: Adjusted Hierarchical Linear Regression Results for % Resident Time in Ambulatory Care Settings/year and Primary Care GME Programs and their Sponsoring Institutions' Characteristics

	Characteristic	Model 1 Coefficient (SE)	Model 2 Coefficient (SE)	Model 3 Coefficient (SE)
Internal Factors	Age of program	-0.02 (0.03)	-0.02 (0.03)	-
	Program type is family medicine	0.4 (1.1)	0.3 (1.0)	-
	# Faculty at Program	**0.03 (0.01) *****	**0.02 (0.01) *****	**-**
	# Residents at Program	-0.02 (0.03)	-0.03 (0.03)	-
	Maximum duty hours/day for PGY1	-0.1 (0.1)	-0.1 (0.1)	-
	# of teaching sites	0.1 (0.2)	0.2 (0.2)	-
External Factors	SI Ownership Private/Nonprofit	-1.3 (1.0)	-	-1.1 (1.0)
	SI in Academic Health Center	1.0 (1.2)	-	1.5 (1.2)
	Hospital beds at SI's primary teaching hospital	0.001 (0.0003)	-	0.001 (0.0003)
	Competing primary care programs at same SI	0.5 (0.4)	-	0.4 (0.4)
	Competing specialty care programs at same SI	-0.04 (0.02)	-	-0.02 (0.02)
	Total CME Medicare GME Funding Received	0.00000007 (0.00000001)	-	0.00000003 (0.00000001)
	SI received Teaching Health Center funding	**6.6 (2.7) ****	**-**	**7.2 (2.7) ****
	Program with ACGME accreditation warnings	**4.8 (2.5) ***	**-**	**4.8 (2.5) ***
	Distance between SI and program (0.4 – 6 miles)	-1.2 (1.1)	-	-1.6 (1.1)
	Distance between SI and program (6 – 37 miles)	-3.1 (1.5)	-	-3.1 (1.4)
	Distance between SI and program (37 – 96 miles)	-0.1 (2.4)	-	-1.0 (2.4)
	Distance between SI and program (96+ miles)	0.1 (2.3)	-	-0.4 (2.3)
County Characteristics (Control)	**% Medicare enrollees in SI's county**	**-0.5 (0.2) ****	**-0.3 (0.1) ****	**-0.5 (0.2) ****
	SI's county is non-metropolitan (rural)	1.9 (2.4)	1.7 (2.3)	2.9 (2.3)
	Health Prof Shortage Area designation for all SI's county	0.3 (0.9)	0.2 (0.1)	0.05 (1.0)
	% under FPL in SI's county	0.1(0.1)	0.1(0.1)	0.1(0.1)
	% Uninsured in SI's county	-0.02 (0.1)	0.02 (0.1)	0.001 (0.1)
	Constant	33.0 (4.3)	30.2 (4.1)	31.9 (3.5)
	$\sqrt{\psi}$	4.0 **	4.7**	3.7 *
	$\sqrt{\theta}$	11.8 **	11.8 **	12.0 *
	Intraclass Correlation (R_2)	0.10	0.14	0.09

**p<.05, **p<.01, ***p<.001*

Figure 3.1: U.S. Map of ACGME-accredited Family Medicine and Internal Medicine Programs & their Sponsoring Institutions, indicating Training Time > 25%

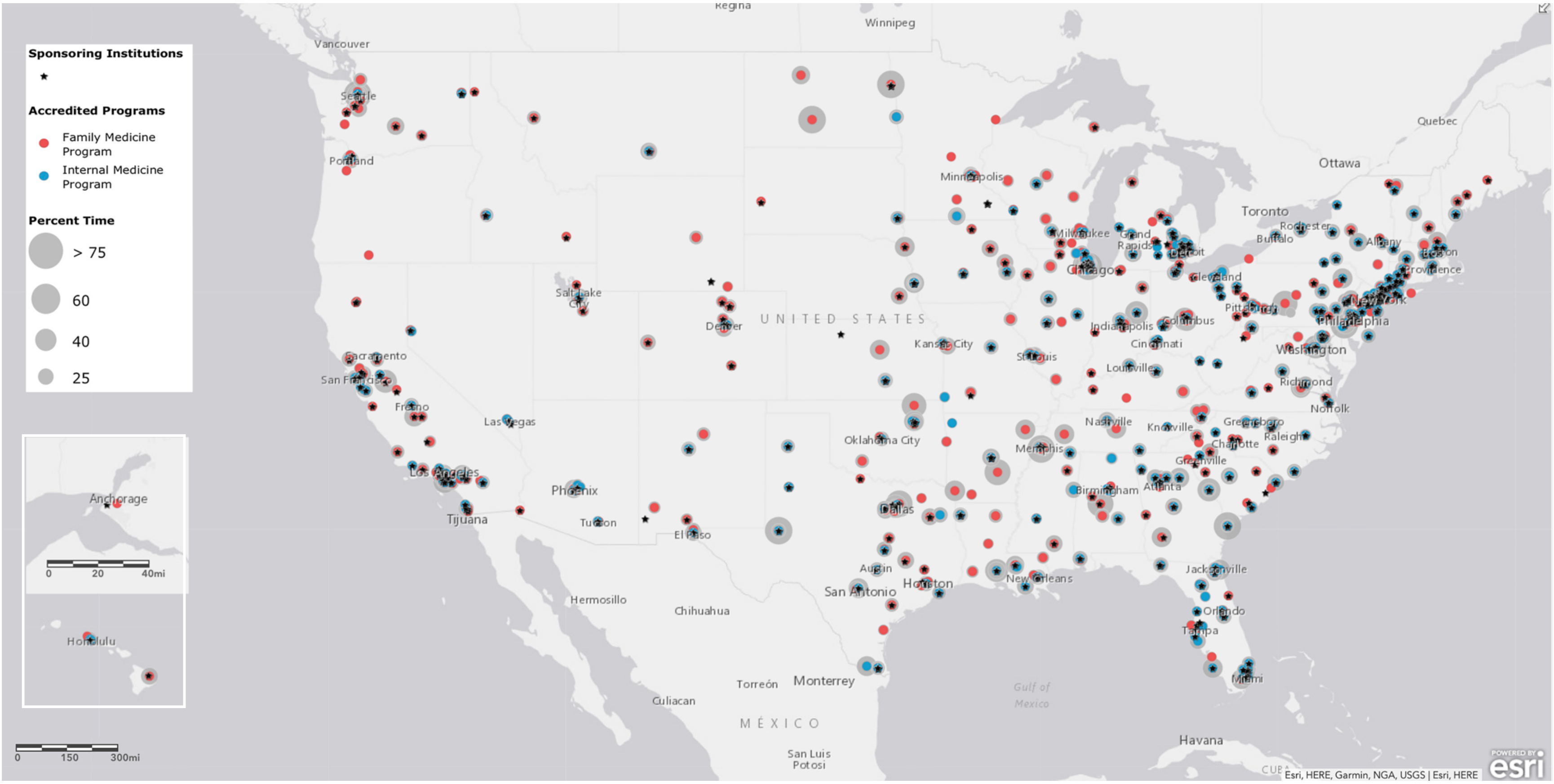
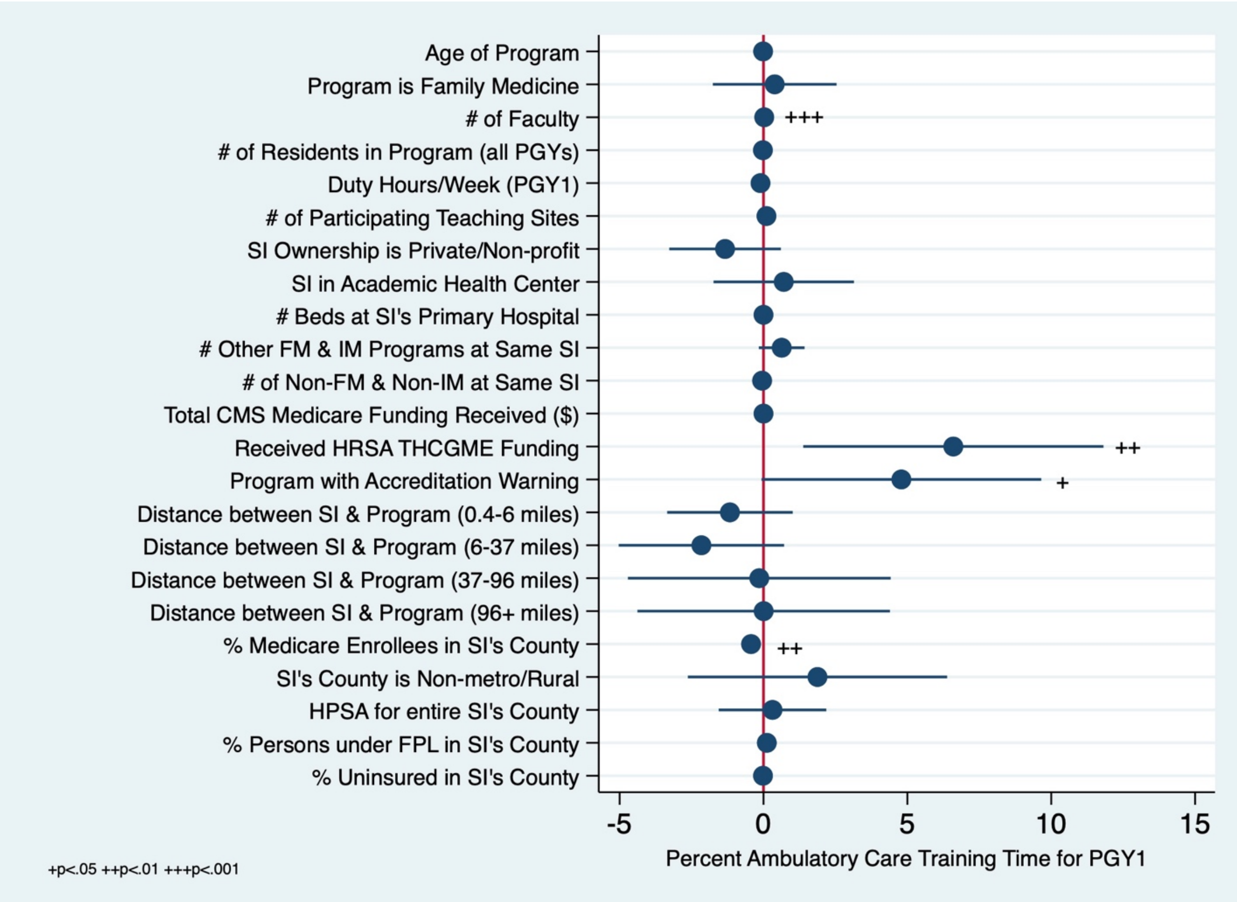


Figure 3.2: Adjusted Coefficient and 95% Confidence Interval Plot for GME Organization and Environmental Factors and Percent Ambulatory Care Time for PGY1s in Family and Internal Medicine



CHAPTER 4: AIM 2
Organizational Factors Associated with Improved Core Competencies among Family Medicine and Internal Medicine Residents

Background:
Competency-based medical education (CBME) is an outcome-based approach to improve physician learning and accountability to patients and health care delivery systems.[78,79] In Graduate Medical Education (GME), CBME was implemented in part by the National Academy of Medicine (NAM) reports "To Err is Human" and "Crossing the Quality Chasm," which illuminated significant patient safety and quality of care deficiencies in the resident learning environment. The NAM concluded in these reports that GME could not systematically assess if resident physicians were fit for practice.[40,80] In response, the Accreditation Council of GME (ACGME) required all GME programs to assess their residents' cumulative learning over the course of their training to be deemed fit for unsupervised practice.[81] To systematically assess GME residents, the ACGME outlines six core clinical competencies, measured by developmental steps toward unsupervised practice via learning rates.[18] Also known as the "Milestones Project," all GME programs are required to assess their residents' learning rates for program accreditation. Upon achievement of the core competencies, a resident is able to graduate from formal training and eligible to become board-certified professionals.[18]

Previous studies argue that physician professional competency is based on both individual learning and a function of their organizational environment.[82] Yet GME's current knowledge on how both residents and their programs learn hinge on sparse studies.[57,83] The lack of GME national-based studies is further perpetuated by the structural environment to which GME operates.[3] As a "cottage industry," GME does not function as a national system, is comprised of independent actors without accountability, and who make autonomous decisions.[3,4] This heterogeneity of organizational factors and until the Milestones Project, have challenged the study of residency program outcomes.[3,11] Of the few studies that examined organizational learning's role for physician learning, organizational level differences account for a high proportion of variation in physicians' procedural skills, and subsequently impact their future practice patterns and behaviors.[38,57,84] However to date, no national study has examined the organizational and external environmental characteristics' impact on family medicine and internal medicine residents' professional competencies while they are still in training.

By drawing on the organizational behavior literature, we posit that organizational factors including differentiation and time of training experiences, size and life cycle of programs, and environmental factors could impact primary care residents' achievement of their core competencies and to some degree, create competency traps, where residency programs may be associated with maladaptive learning curves.[11,28,30,35,85] This study examines the internal and external organizational factors associated with the ACGME core clinical competencies. By examining influences from internal and external organizational factors, this study aims to identify the predictors that impact GME program learning rates for residents training in ACGME-accredited Family Medicine and Internal Medicine programs.

Methods

Sample/Dataset:

A longitudinal multi-level dataset was analyzed for all U.S. accredited-Family Medicine and Internal Medicine residency programs who reported their residents' clinical competency scores from 2014-17 and linked at the program level to the national ACGME Milestones Project. Data from 784 programs (representing 10,862 residents assessed over three years) were linked to the 2016-2017 American Medical Association (AMA) and Association of American Medical Colleges (AAMC)'s National GME Census of ACGME-accredited programs. The GME Census survey is completed annually by all U.S. GME programs with a 95.7% response rate.[61] The data was also linked to each program's primary hospital training site to the 2016 CMS Cost Reports, and each GME sponsoring institution's Federal Information Processing Standards (FIPS) county codes to the 2015-2016 Area Health Resource File (AHRF). [58,59]

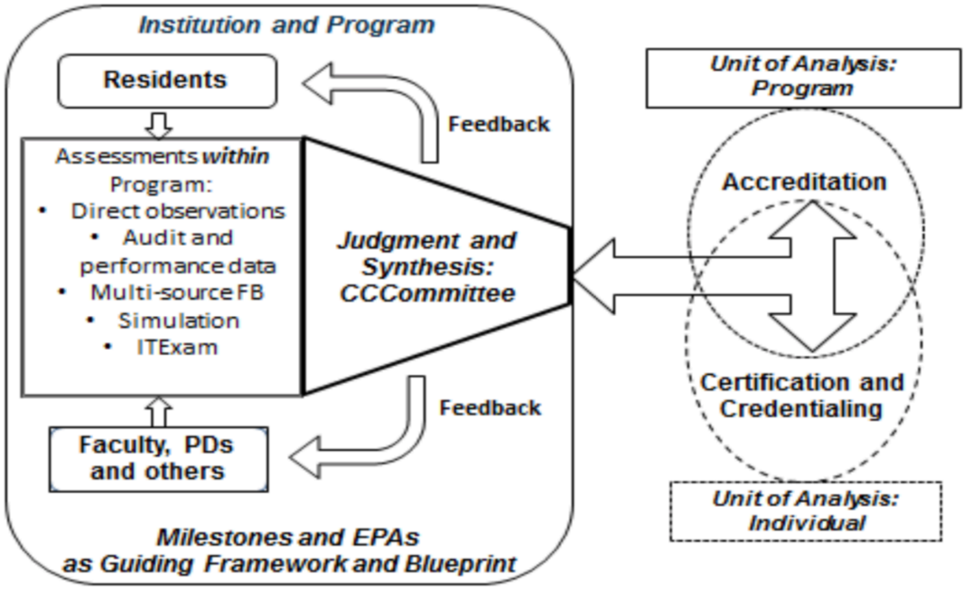
Main Outcome:

Resident program learning rates were calculated based on the average of individual resident level clinical competency Milestones scores for each program. For this analysis, three of the six core ACGME clinical competencies were selected for our main outcomes: patient care, practice-based learning and improvement, and systems-based practice.[18,86] These core competencies were selected due to their measurement constructs that include organizational level learning. The specifications for the selected core competencies are as follows[18]:

- **Patient Care:** Residents must be able to provide patient care that is compassionate, appropriate, and effective for the treatment of health problems and the promotion of health.
- **Practice Based Learning and Improvement:** Residents must demonstrate the ability to investigate and evaluate their care of patients, to appraise and assimilate scientific evidence, and to continuously improve patient care based on constant self-evaluation and life-long learning.
- **Systems-Based Practice:** Residents must demonstrate an awareness of and responsiveness to the larger context and system of health care, as well as the ability to call effectively on other resources in the system to provide optimal health care.

Figure 4.1 describes the process to which items are assessed and core competencies scored by program faculty and subsequently reported to the ACGME. See "Judgement and Synthesis" via the Clinical Competency Committee (CCC).[18]

Figure 4.1: Systematic Assessment of Residents via the ACGME Milestones Project



Every six months and over 3 years (6 total assessments/resident), each programs' faculty rater in their clinical competency committee (CCC) assign an item score to the resident's core competency, ranging from a score of 0-9 over 5 levels: Level 0 (Score 0): Assumes resident is not ready for program; Level 1 (Score 1-2): Assumes resident has some prior training in medicine (i.e.: medical school); Level 2 (Score 3-4): Resident demonstrates developing competency and is advancing in program; Level 3 (Score 5-6): Resident demonstrates developing competency and is near completion of program; Level 4 (Score 7-8): Resident has demonstrated achieving competency and is fit for unsupervised practice; Level 5 (Score 9): Resident has performed beyond the performance expected. At end of the resident's training period, each resident will have been rated 6 times. These longitudinal Milestone scores for each resident were then collapsed at the program level at each time point and used to calculated the program level learning rate.

Predictor Variables:
Organizational factors were categorized as internal and external to the program, and include control variables at the county level.

Internal Organizational Factors
Internal organizational characteristics for residency programs included: differentiation of training characteristics (primary care specialty: family medicine or internal medicine, total number of teaching sites); size of programs (number of faculty and residents in program); and time factors in the learning environment (residents' maximum duty hours/day; resident % time in the ambulatory care setting).

External Factors
Factors external to the residency program included characteristics of the sponsoring institution that sponsors each program: if sponsoring institutions' ownership type was private/non-profit, set

within an academic health center, number of hospital beds, age (life cycle) of program sponsored, number of (competing) primary care and non-primary care residency programs; total amount of CMS Medicare GME funding received in 2016; and receipt of federal Teaching Health Center GME funding. Accreditation warnings from the ACGME were also modeled. Structural characteristics including distance in miles between the sponsoring institution and the residency program's primary address listed in the GME Census were calculated. Addresses and geocoordinates were used for sponsoring institutions and their accredited programs' primary teaching site, using the *georoute* command in STATA.[64] Subsequently, this continuous variable for distance of travel was transformed into a categorical variable due to a large positive skew distribution, and divided into values by 50%ile, 75%ile, 90%ile, and the 95th %ile of its distribution.

Control Variables

We also modeled county-level antecedent variables that adjusted for sponsoring institution's county characteristics and could impact the internal and external characteristics of the program. These include: whether the entire county was designated as a health professions shortage area (HPSA); the proportion of the county population living below the Federal Poverty Level (FPL); the county's level of rurality using the AHRF's Urban-Rural Continuum Codes; the proportion of the county population that is enrolled in Medicare; and the proportion of the county population that is uninsured.

Analytic Approach:

Internal characteristics and external factors, plus county-level controls for family medicine and internal medicine residency programs were compared using chi-squared for categorical independent variables and t-tests for continuous independent variables.

To estimate the relationship of organizational factor predictors and Milestone scores, we used hierarchical growth curve models fitted for each of the three main outcomes (the three core competencies listed above) with internal and external organizational factors as predictor variables. The model also incorporated the control variables listed above. Growth curve modeling was selected to estimate inter-unit level differences across clustered paneled data, which would account for programs clustered in sponsoring institutions. This analytic approach is commonly used for education learning rate studies that model the shape and variance of student learning trajectories over time.[65,87] Due to the nonlinear growth observed in our unadjusted analyses, timepoints were transformed into second-order polynomials to estimate smoother linear curve rates.[65] Random effects were used to account for variance of each GME program against the overall GME program mean scores. We fitted three analytic models (internal factors, internal with external, and internal/external plus county level controls) for each of the three core competencies selected (milestones1-3_{ij}):

$$milestones\_n_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_1 t_{ij}^2 + \beta_1 X_j + u_{j0} + u_{j1} t_{ij} + \epsilon_{ij};$$

Model 1:

$\beta_1 t_{ij}$ = t_{ij} = time of Milestone assessment; $\beta_1 t_{ij}^2$ = t^2_{ij} = quadratic function of time of Milestone assessment ; βX_j = {x_{1j} = outpatienttime$_j$; x_{2j} = rotations$_j$; x_{3j} = dutyhours$_j$; x_{4j} = programsize$_j$; x_{5j} = faculty$_j$; x_{6j} = programtype$_j$} ; u_{j0} = random level effects intercept and $u_{j1} t_{ij}$ = random level time slope that deviates from the overall mean linear growth rate. $\epsilon_{ij} \sim N(0, \theta)$, where ϵ_{ij} is the residual of the program nested and assumed to be normally distributed.

Model 2:
$\beta_1 t_{ij}$ = t_{ij} = time of Milestone assessment; $\beta_1 t_{ij}^2$ = t^2_{ij} = quadratic function of time of Milestone assessment ; βX_j = {x_{1j} = outpatienttime$_j$; x_{2j} = rotations$_j$; x_{3j} = dutyhours$_j$; x_{4j} = programsize$_j$; x_{5j} = faculty$_j$; x_{6j} = programtype$_j$; x_{7j} = miles$_j$; x_{8j} = age$_j$; x_{9j} = academichealthcenter$_j$; x_{10j} = competingprograms$_j$; x_{11j} = specialtyprograms$_j$; x_{12j} = ownership$_j$; x_{13j} = beds$_j$; x_{14j} = accreditationstatus$_j$} ; u_{j0} = random level effects intercept and $u_{j1}t_{ij}$ = random level time slope that deviates from the overall mean linear growth rate. $\epsilon_{ij} \sim N(0, \theta)$, where ϵ_{ij} is the residual of the program nested and assumed to be normally distributed.

$$milestones\_n_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_1 t_{ij}^2 + \beta_1 X_j + \beta_1 Z_j + u_{j0} + u_{j1} t_{ij} + \epsilon_{ij};$$

Model 3:
$\beta_1 t_{ij}$ = t_{ij} = time of Milestone assessment; $\beta_1 t_{ij}^2$ = t^2_{ij} = quadratic function of time of Milestone assessment ; βX_j = {x_{1j} = outpatienttime$_j$; x_{2j} = rotations$_j$; x_{3j} = dutyhours$_j$; x_{4j} = programsize$_j$; x_{5j} = faculty$_j$; x_{6j} = programtype$_j$; x_{7j} = miles$_j$; x_{8j} = age$_j$; x_{9j} = academichealthcenter$_j$; x_{10j} = competingprograms$_j$; x_{11j} = specialtyprograms$_j$; x_{12j} = ownership$_j$; x_{13j} = beds$_j$; x_{14j} = accreditationstatus$_j$; x_{15j} = cms_\$$_j$; x_{16j} = thc$_j$ }; βZ_j = {z_{1j} = hpsa$_j$; z_{2j} = medicare$_j$; z_{3j} = uninsured$_j$; z_{4j} = poverty$_j$; z_{5j} = rural$_j$} ; u_{j0} = random level effects intercept and $u_{j1}t_{ij}$ = random level time slope that deviates from the overall mean linear growth rate. $\epsilon_{ij} \sim N(0, \theta)$, where ϵ_{ij} is the residual of the program nested and assumed to be normally distributed.

The three models were tested for multicollinearity using variance inflation factors (VIF), which reported a mean VIF of 3.3.[66] Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were calculated as a post-test estimate to determine model fit.[65] Sensitivity analyses were conducted to assess the robustness of our results including models without clustering using multivariate linear regression and also a hierarchical linear model using only county-level variables to examine the association of the antecedent macro county-level variables. Results with $p<.05$ were considered statistically significant. Analyses were performed using STATA 15.1.[67] This study received a human subjects exemption from the University of California, Berkeley Institutional Review Board (IRB).

Results
Table 4.1 compares family medicine and internal medicine residency programs by internal organizational characteristics, external factors, and county control variables. Descriptive findings between the two types of primary care medical specialties were similar to those found in Aim 1.

Table 4.2 describes the baseline differences of Milestone scores observed between primary care specialties, with internal medicine programs reporting higher baseline Milestone scores than those in family medicine. For baseline patient care scores, internal medicine programs scored 1.0 higher than the overall mean versus family medicine which scored 0.8 lower; in practice-based learning and improvement scores, internal medicine programs had even higher baseline scores from the overall mean versus family medicine (1.2 higher vs. 1.0 lower); and for systems-based practice scores, internal medicine programs also had scores 1.0 higher than the overall mean compared to family medicine which reported scores 1.0 lower.

Improvement of Milestone scores between baseline and graduation also found internal medicine programs with higher scores than family medicine programs. For patient care scores, internal medicine programs had an average (SD) Milestone score of 7.2 (0.4) compared to 6.9 (0.6) for family medicine; in practice-based learning and improvement scores, internal medicine programs had an average (SD) Milestone score of 7.1 (0.4) compared to 6.5 (0.7) for family medicine; and for systems-based practice scores, internal medicine programs had an average (SD) Milestone score of 7.2 (0.4) compared to 6.7 (0.7) for family medicine.

Tables 4.3 to 4.5 report the adjusted analysis of learning rates, with mean clinical competency score improvement (95% CI) reported for every biannual timepoint in patient care, practice-based learning and improvement, and systems-based practice competencies: 0.98 (0.92 to 1.03), 0.89 (0.84 to 0.97), 0.91 (0.86 to 0.96) improvement at each time point, respectively ($p<.001$).

The learning rate trajectories between 2014-2017 indicated only external factors were associated with higher learning rates and only for the patient care and practice-based learning and improvement core competencies: Patient care: programs located with 6-37 miles away from their sponsoring institution (0.13, 0.0089 to 0.25, $p<.05$) and those in rural counties (0.23, -0.0010 to 0.47, $p<.05$); Practice based learning and improvement: programs in counties with higher % of those living under the Federal Poverty Level (0.011, 0.0016 to 0.021, $p<.05$).

For factors negatively associated with learning rate trajectories, external factors and county control variables were statistically significant for all three core competencies examined. Additionally, one internal factor, programs reporting greater than the mean maximum duty hours/day was found to be statistically significant. In the patient care learning rates, programs cited with accreditation warnings (-0.26, -0.49 to -0.032, $p<.05$) and those set in counties designated as a health professional shortage area (-0.13, -0.21 to - 0.047, $p<.01$); for practice based learning and improvement learning rates: programs in counties with higher % of uninsured (-0.011, -0.021 to -0.0039, $p<.05$), designated as a health professional shortage area (-0.13, -0.22 to - 0.038, $p<.01$); and for systems-based practice learning rates: programs who reported greater than the mean maximum duty hours/day (-0.17, -0.36 to -0.0030, $p<.05$); in counties with higher % of Medicare enrollees (-1.27, -2.61 to -0.026, $p<.05$) higher % of uninsured (-0.0094, -0.019 to -0.00022, $p<.05$) and designated as a health professional shortage area (-0.12, -0.021 to -0.0039, $p<.01$).

Figures 4.2 illustrates boxplots for Milestone score distributions by family medicine and internal medicine and by core competencies. The distributions for internal medicine programs had a narrower spread than family medicine, with the latter reporting more outliers for the patient care competency. For practice-based learning and improvement and systems-based practice, family medicine residency programs include more positive outliers compared to internal medicine residency programs.

Figures 4.3 and 4.4 illustrate the unadjusted and adjusted learning rates for each core competency examined. In Figure 4.4, learning rates were higher for internal medicine programs.

Additionally, for patient care and practice-based learning and improvement core competencies, some family medicine programs were found outside than the majority of programs fitted within the growth curve model, indicating deviation.

Discussion

Compared to internal program factors, external factors were more positively associated with improved 3-year program learning rates after adjusting for county level characteristics. For organizational factors found negatively associated with higher learning rates, external factors including accreditation warnings were found to impact Milestone scores for the patient care core competency. GME program structural arrangements and geographic setting also were associated to impact program learning rates. Programs further than 6 miles away (but less than 37 miles) from their sponsoring institution and primary hospital setting were associated higher learning rates for the patient care competency. For programs in rural counties, and in counties with a higher % of residents living under the FPL, learning rates were higher for the core competencies in patient care and practice-based learning and improvement, respectively. On the contrary, programs in health professional shortage areas and in counties with higher % of Medicare enrollees and uninsured were negatively associated with higher learning rates in practice-based learning and improvement and systems-based practice competencies. The strongest predictors for improving learning rates were associated with accumulation of Milestones assessments over time and for GME programs accredited in internal medicine, as indicated by the increasing scores over time and differences between the two specialty's learning curve rates.

The differences between family medicine and internal medicine residency program learning rates are noteworthy. Internal medicine residency programs' organizational level learning responded more positively to their organizational and environmental factors, as inferred by the higher learning rates in patient care, practice-based learning and improvement, and systems-based practice. These factors could be attributed to the structural differences between family medicine and internal medicine programs. Whereas in family medicine, programs are often in outpatient settings, versus internal medicine programs which are primarily in hospital-based settings and subject to more accreditation and compliance oversight related to patient safety and quality improvement efforts.[62,88,89] The awareness of these efforts could be impacting the way programs are rating the core competencies in patient care, practice-based learning and improvement, and systems-based practice, which incorporate organizational level interventions that prioritize patient safety. Our findings with external organizational factors as better predictors for learning rates may imply that higher presence of external oversight may influence the core competencies, however given the available data analyzed, our models did not account for organizational learning interventions specific to hospital-based settings.

Similarly, our findings for external forces as better predictors also apply to programs cited with accreditation warnings, which were found to have slower learning rates for the patient care competency. Programs cited for violations by ACGME may not be providing a learning environment for residents to meet this core competency that assesses for improvement in patient outcomes. This could be explained by one of the most frequent accreditation warnings issued to family medicine programs, insufficient number of patients cared for in continuity clinics.[73] Or simply, programs that receive ACGME warnings may be providing poor training quality and thus inadequate patient encounters impact residents' ability to achieve competence in patient care. As reported in Aim 1, further study is needed to understand the association of types of accreditation citations received, and how citations play a role in impacting training design for residents.[73]

GME program structural design via geospatial characteristics were found to be inconsistent but statistically significant predictors. The improved Milestone scores in the patient care competency for programs located between 6 miles and 37 miles away from their sponsoring institution and primary hospital training site, set in rural counties, and in counties with a higher % of residents living under the FPL indicate that residents training in outpatient primary care settings (vs. inpatient) achieve greater learning gains for certain core competencies. On the contrary, slower learning rates for practice-based learning and improvement and systems-based practice competencies associated with residents in health professional shortage areas, and in counties with higher Medicare and uninsured patients could be attributed to demanding practice settings impacting the programs set in these counties.[90] For programs set in counties with a shortage of primary care physicians and who serve a higher proportion of Medicare and uninsured patients, our findings suggest that GME program learning rates are impacted by the external county environment, and ultimately making their way into the resident learning space.

The findings for systems-based practice competencies are especially important to highlight given the increasing interest to understand the role of health systems in GME.[91–94] Greater than the mean maximum duty hours/day is associated with slower learning rates in systems-based practice may suggest that residents struggling to meet the demands of their delivery system are adversely impacted by increased work hours. Limits on resident work hours remain contentious in GME policy.[95–97] Time constraints during resident training highlight the competing demands of resident service with patient care versus time needed to master a competency for individual learning needs. As described in other studies, GME is incentivized to prioritize hospital and inpatient-delivery of care.[4,11] Service demands through increased work hours across multiple service lines may be pulling residents from their core learning needs to understand the requirements to meet the competency of systems-based practice and demonstrating competence to their residency program. For example, if residents are pulled into non program-related service lines, opportunities to learn systems-based practice via quality improvement projects in their programs become less available to them.[88] Further study is needed to understand if residents immersed more in systems-level learning opportunities could achieve faster competency in systems-based practice.

Limitations

The study's limitations should be considered when interpreting our findings. First, the data sources are assessments subject to social desirability bias and other faculty rater biases. This Milestones data represents the first cohort evaluated under this new ACGME policy change. This study could not compare Milestones scores across different cohort years to examine validity and reliability although previous studies on each of the specialties have described efforts to understand the Milestone scores.[86,98–100] Secondly, the medical specialties selected are not representative of the entire primary care training workforce across different practice settings, including osteopathic residents and nurse practitioners, and whether the Milestones adequately assess for primary care competencies.[61,62] Lastly, assigning random effects at the program level in our analytic models may not explicitly account for all factors, including organization level learning interventions that may be associated with resident learning rates, thus unmeasured confounding cannot be ruled out. Also, Milestone scores are assigned by faculty rater groups and these groups' characteristics were not available in the data, so we are unable to account these factors into our analytic models. Future studies for the Milestones Project could include modeling directly observed behaviors of faculty and utilizing Rasch rater harshness analysis to further account for the variation of learning rates.

Conclusion:

Improved resident learning rates in the ACGME Milestones for family medicine and internal medicine residency programs between 2014-2017 were associated more with external factors than internal factors. Learning rates across the core competencies in patient care, practice-based learning and improvement, and systems-based practice may be dependent on the program's geographic setting, structural characteristics, and the type of the learning experience for residents. These findings could inform policy discussions on how the external factors ultimately shape resident learning experiences and subsequently impact the learning rates of residents.

Table 4.1: Characteristics of Primary Care GME Programs

Factor	Characteristic	Family Medicine (n=427)	Internal Medicine (n=357)	Overall (n=784)
Internal	# Residents at Program (mean, SD)	22.9 (0.4)	63.9 (2.0) ***	41.6 (1.2)
	# of Faculty at Program (mean, SD)	13.3 (0.7)	103.0 (7.3) ***	54.2 (3.7)
	Maximum Duty Hours/Week for PGY1 (mean, SD)	16.1 (0.1)	16.4 (0.3)	16.2 (0.1)
	% ambulatory care time for PGY1 (mean, SD)	25.4 (0.7)	25.7 (0.5)	25.5 (0.5)
	# of teaching sites (mean, SD)	3.4 (0.1) ***	2.7 (0.1)	3.1 (0.9)
External	Distance in miles between Program's Teaching Site and Sponsor (mean, SD)	22.0 (4.1)	14.3 (6.2)	18.5 (3.6)
	Program in Academic Health Center (n, %)	126 (29.1%)	161 (45.1%) ***	287 (36.1%)
	Hospital beds at program's primary teaching hospital (mean, SD)	469 (13.7)	532 (17.5)	498 (11.0)
	# of Competing Primary Care Program at Same SI (mean, SD)	1.5 (0.1)	1.3 (0.1)	1.4 (0.6)
	# of Non-Primary Care Programs at Same SI (mean, SD)	22.8 (1.6)	32.9 (2.0) ***	27.4 (1.3)
	Age of Program (mean, SD)	33.8 (0.6)	48.2 (0.9) ***	40.4 (0.6)
	Program's SI Ownership Type (n, %) Private Non-profit	219 (51.3%)	208 (58.3%)	427 (54.5%)
	Total CMS Medicare GME funding received (mean, SD)	$6,196,968 ($390,702)	$9,727,616*** ($587,584)	$7,797,316 ($347,362)
	SI received Teaching Health Center funding (n, %)	20 (4.6%) **	4 (1.1%)	24 (3.0%)
	Program with ACGME accreditation warnings (n, %)	20 (4.6%) *	6 (1.7%)	26 (3.3%)
County Characteristics	% Medicare Enrollees in SI's county (mean, SD)	16.8 (0.0)*	16.2 (0.0)**	16.5 (0.0)
	SI's county is non-metropolitan/rural (n, %)	24 (5.6%) ***	5 (1.4%)	29 (3.7%)
	Health Professional Shortage Area designation for all SI's county (n, %)	171 (40.0%)	134 (37.5%)	306 (39.0%)
	% under FPL in SI's county (mean, SD)	15.9 (0.2)	16.2 (0.3)	16.0 (0.2)
	Uninsured % in SI's county (mean, SD)	16.4 (0.2)*	15.6 (0.3)	16.1(0.2)
	Census Region (n, %)			
	West	96 (22.5%)***	51 (14.3%)	147 (18.8%)
	Midwest	121 (28.3%)	84 (23.8%)	205 (26.1%)
	South	136 (31.9%)	94 (26.3%)	230 (29.3%)
	Northeast	74 (17.3%)	128 (35.9%) ***	202 (25.8%)

**p<.05, **p<.01, ***p<.001*

PGY1 = first year resident; SI = sponsoring institution; CMS = Centers for Medicaid and Medicare Services; ACGME = Accreditation Council for Graduate Medical Education; FPL = Federal Poverty Level

Table 4.2: Mean Clinical Competency Scores for Primary Care GME Programs by Timepoint

Core Clinical Competency	Family Medicine (n=427)	Internal Medicine (n=357)	Overall (n=784)
Patient Care (mean, SD)			
Timepoint 1	2.1 (0.8)	3.9 (0.8)***	2.9 (1.2)
Timepoint 2	3.2 (0.8)	4.6 (0.4)***	3.8 (1.0)
Timepoint 3	4.4 (0.8)	5.3 (0.6)***	4.8 (0.9)
Timepoint 4	5.2 (0.8)	5.8 (0.6)***	5.4 (0.7)
Timepoint 5	6.2 (0.7)	6.5 (0.5)***	6.3 (0.2)
Timepoint 6	6.9 (0.6)	7.2 (0.4)***	7.0 (0.5)
Practice-based Learning & Improvement (mean, SD)			
Timepoint 1	1.9 (0.9)	4.1 (1.0)***	2.9 (1.4)
Timepoint 2	2.9 (0.9)	4.8 (0.8)***	3.8 (1.3)
Timepoint 3	4.1 (0.8)	5.4 (0.7)***	4.7 (1.0)
Timepoint 4	4.8 (0.7)	5.8 (0.7)***	5.3 (0.9)
Timepoint 5	5.7 (0.7)	6.5 (0.7)***	6.1 (0.8)
Timepoint 6	6.5 (0.7)	7.1 (0.4)***	6.8 (0.7)
Systems-based Practice (mean, SD)			
Timepoint 1	2.1 (0.9)	4.1 (0.9)***	3.1 (1.3)
Timepoint 2	3.2 (0.8)	4.7 (0.8)***	3.9 (1.1)
Timepoint 3	4.3 (0.8)	5.4 (0.7)***	4.8 (0.9)
Timepoint 4	5.0 (0.7)	5.8 (0.6)***	5.4 (0.8)
Timepoint 5	6.0 (0.7)	6.5 (0.6)***	6.3 (0.7)
Timepoint 6	6.7 (0.7)	7.2 (0.4)***	6.9 (0.2)

**p*<.05, ***p*<.01, ****p*<.001

Figure 4.2: Distribution of Clinical Competency Scores for Primary Care GME Programs by ACGME Core Competency

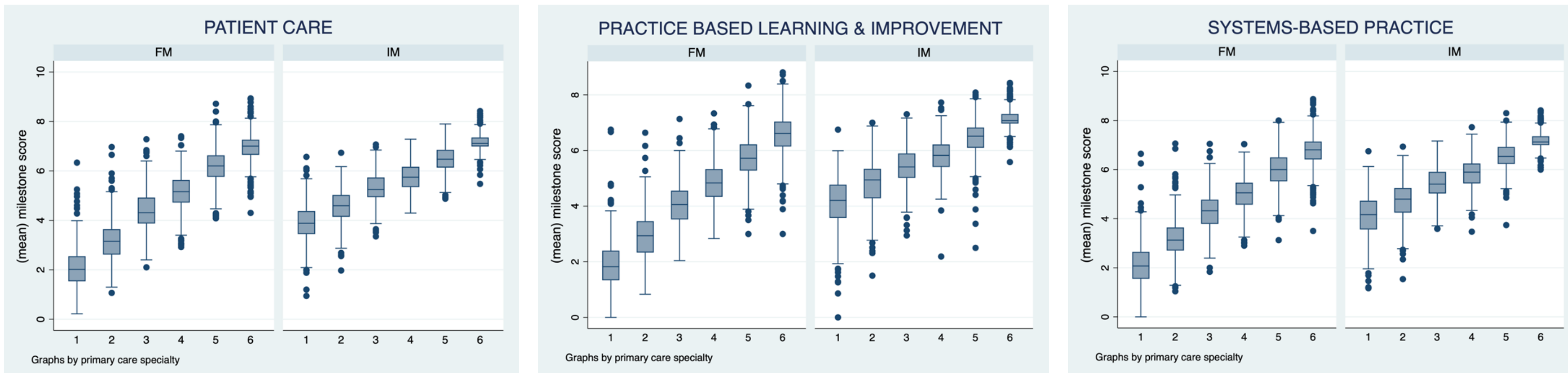


Table 4.3: Adjusted Hierarchical Linear Regression Models for Clinical Core Competency Scores and Primary Care GME Organizational Factors:

PATIENT CARE COMPETENCY

Factor	Predictor	Model 1 Coefficient (95% CI)	Model 2 Coefficient (95% CI)	Model 3 Coefficient (95% CI)
Internal Factors	Assessment timepoint	0.98 (0.93- 1.03)***	0.98 (0.930 -1.03)***	0.98 (0.93 - 1.03)***
	Assessment timepoint^2	-0.023 (-0.028 - -0.017)***	-0.023 (-0.028 - -0.017)***	-0.023 (-0.028 - -0.017)***
	Program type is family medicine	-0.56 (-0.65- -0.47)***	-0.56 (-0.66 - -0.47)***	-0.56 (-0.65 - -0.47)***
	# Residents at Program	-0.00074 (-0.0027 - 0.0012)	-0.00072 (-0.0028 -0.0013)	-.0004 (-0.0024 - 0.0016)
	# Faculty at Program	0.00056 (0.0000075 - 0.0011)*	0.00054 (-0.000052 -0.0011)	0.00044 (-0.00018 -0.0011)
	>mean duty hours/day for PGY1 (16 hours/day)	-0.152 (-0.323 - 0.0193)	-0.16 (-0.34 - 0.018)	-0.16 (-0.34 - 0.018)
	% ambulatory care time for PGY1	-0.00042 (-0.0038 - 0.0029)	0.000063 (-0.0032 - 0.0034)	-0.00025 (-0.0036 -0.0031)
	# of teaching sites	0.0097 (-0.0061- 0.026)	0.0078 (-0.0075 -0.023)	0.0083 (-0.0069 -0.024)
External Factors	Distance between program and SI (0.4 – 6 miles)		0.053 (-0.047 - 0.15)	0.037 (-0.062 - 0.14)
	Distance between program and SI (6 – 37 miles)		0.14 (0.015 - 0.27)*	0.13 (0.0089 -0.25)*
	Distance between program and SI (37 – 96 miles)		0.16 (-0.11 - 0.43)	0.12 (-0.14 - 0.38)
	Distance between program and SI (96+ miles)		-0.11 (-0.33 - 0.12)	-0.15 (-0.38 - 0.086)
	SI in Academic Health Center		0.082 (-0.022 - - 0.19)	0.039 (-0.062 -0.14)
	Hospital beds at SI's primary teaching hospital (standardized)		0.014 (-0.041 - - 0.068)	0.016 (-0.039 -0.071)
	Competing primary care programs at same SI		0.023 (-0.072 - 0.12)	0.049 (-.045 - 0.14)
	Competing specialty care programs at same SI (standardized)		-0.032 (-0.10 - 0.040)	-0.021 (-0.093 - 0.050)
	Age of program (standardized)		0.0096 (-0.035 - 0.054)	0.0044 (-0.040 - 0.049)
	SI Ownership Private/Nonprofit		-0.0067 (-0.087 - -0.073)	-0.0099 (-0.089 - 0.069)
	Total CME Medicare GME Funding Received (standardized)		-0.0057 (-0.069 - 0.059)	-0.025 (-0.0930 - 0.043)
	SI received Teaching Health Center funding		0.010 (-0.21 - 0.25)	0.058 (-0.17 - 0.29)
	Program with ACGME accreditation warnings		-0.26 (-0.48 - -0.035)*	-0.26 (-0.49 - -0.032)*

County Characteristic	% Medicare enrollees in SI's county			-1.27 (-2.57 - 0.055)
	SI's county is non-metropolitan (rural)			0.23 (-0.0010 - 0.47)*
	Health Prof Shortage Area designation for all SI's county			-0.13 (-0.21 - - 0.047)**
	% under FPL in SI's county			0.0078 (-0.0012 - 0.016)
	% Uninsured in SI's county			-0.0083 (-0.017 - 0.00079)
Constant		2.26 (2.10 - 2.42)	1.90 (1.75 - 2.04)	2.45 (2.10 - 2.79)
$\sqrt{\psi}$		1.10 (1.02 - 1.19)	1.09 (1.01 - 1.19)	1.09 (1.01 - 1.18)
$\sqrt{\theta}$		0.37 (0.36 -0.39)	0.37 (0.36 - 0.39)	0.37 (0.36 - 0.39)
AIC		7592.88	7599.77	7590.39
BIC		7676.81	7767.63	7760.53

**p*<.05, ***p*<.01, ****p*<.001*

Table 4.4: Adjusted Hierarchical Linear Regression Models for Clinical Core Competency Scores and Primary Care GME Organizational Factors:

PRACTICE-BASED LEARNING & IMPROVEMENT COMPETENCY

Factor	Predictor	Model 1 Coefficient (95% CI)	Model 2 Coefficient (95% CI)	Model 3 Coefficient (95% CI)
Internal Factors	Assessment time point	0.89 (0.84 - 0.97)***	0.89 (0.84 - 0.94)***	0.89 (0.84 - 0.94)***
	Assessment timepoint^2	-0.017 (-0.023 - -0.011)***	-0.017 (-0.023 - -0.011)***	-0.017 (-0.023 - -0.011)***
	Program type is family medicine	-1.03 (-1.13 - -0.93)***	-1.03 (-1.13 - -0.93)***	-1.03 (-1.13 - -0.93)***
	# Residents at Program	-0.0011 (-0.0031 - 0.00095)	-0.0012 (-0.0034 - 0.00099)	-0.0011 (-0.0033 - 0.0010)
	# Faculty at Program	0.00042 (-0.00017 - 0.0010)	0.00027 (-0.00034 -0.00088)	0.00023 (-0.00041- 0.00087)
	>mean duty hours/day for PGY1	-0.11 (-0.28 -0.059)	-0.12 (-0.29 - 0.058)	-0.12 (-0.29 - 0.058)
	% ambulatory care time for PGY1	-0.00079 (-0.0046 - 0.0030)	-0.00039 (-0.0041- 0.0034)	-0.00058 (-.0044 - .0032)
	# of teaching sites	0.0084 (-0.0077 - 0.025)	0.0056 (-0.010 - -0.022)	0.0061 (-0.0095 - 0.022)
External Factors	Distance between program and SI (0.4 – 6 miles)		0.033 (-0.074 0.14)	0.017 (-0.087- 0.12)
	Distance between program and SI (6 – 37 miles)		0.045 (-0.092 -0.18)	0.039 (-0.098 -0.18)
	Distance between program and SI (37 – 96 miles)		0.20 (-0.065 - 0.46)	0.21 (-0.050 - 0.47)
	Distance between program and SI (96+ miles)		-0.15 (-0.38 - 0.072)	-0.16 (-0.39 - 0.075)
	SI in Academic Health Center		0.036 (-0.073 - 0.15)	-0.011 (-0.12 - 0.097)
	Hospital beds at SI's primary teaching hospital (standardized)		0.0057 (-0.058 - 0.070)	0.0096 (-0.057 -0.076)
	Competing primary care programs at same SI		0.066 (-0.037 - 0.17)	0.081 (-0.022 - 0.18)
	Competing specialty care programs at same SI (standardized)		0.016 (-0.056 - 0.086)	0.030 (-0.041 - 0.10)
	Age of program (standardized)		0.017 (-0.032 - 0.067)	0.0094 (-0.041- 0.060)
	SI Ownership Private/Nonprofit		0.021 (-0.067 - 0.11)	0.016 (-0.072 - 0.10)
	Total CME Medicare GME Funding Received (standardized)		-0.010 (-0.078 - 0.058)	-0.039 (-0.11 - 0.033)
	SI received Teaching Health Center funding		0.093 (-0.17 - 0.36)	0.12 (-0.13 - 0.37)
	Program with ACGME accreditation warnings		-0.19 (-0.43 -0.048)	-0.19 (-0.43 - 0.054)

County Characteristics	% Medicare enrollees in SI's county			-0.95 (-2.32 -0.43)
	SI's county is non-metropolitan (rural)			-0.019 (-0.29 - 0.26)
	Health Prof Shortage Area designation for all SI's county			-0.13 (-0.22 - -0.038)**
	% under FPL in SI's county			0.011 (0.0016 - 0.021)*
	% Uninsured in SI's county			-0.011 (-0.021 - -0.00086)*
Constant		2.67 (2.49 - 2.84)	2.59 (2.38 - 2.81)	2.81 (2.42- 3.19)
$\sqrt{\psi}$		1.17 (1.09 - 1.26)	1.16 (1.09 - 1.25)	1.16 (1.08 - 1.25)
$\sqrt{\theta}$		0.42 (0.39 - 0.45)	0.42 (0.39 - 0.45)	0.42 (0.39 - 0.45)
AIC		8644.49	8654.30	8649.12
BIC		8728.43	8822.16	8849.26

**p<.05, **p<.01, ***p<.001*

Table 4.5: Adjusted Hierarchical Linear Regression Models for Clinical Core Competency Scores and Primary Care GME Organizational Factors:

SYSTEMS-BASED PRACTICE COMPETENCY

Factor	Predictors	Model 1 Coefficient (SE)	Model 2 Coefficient (SE)	Model 3 Coefficient (SE)
Internal Factors	Assessment time point	0.91 (0.86 - 0.96)***	0.91 (0.86 - 0.96)***	0.91 (0.86 - 0.96)***
	Assessment timepoint^2	-0.019 (-0.025 - -.0014)***	-0.019 (-0.025 - -.0014) ***	-0.019 (-0.025 - -.0014) ***
	Program type is family medicine	-0.81 (-0.90 - -0.072)***	-0.82 (-0.92 - -0.072) ***	-0.81 (-0.91 - -0.071) ***
	# Residents at Program	-0.0013 (-0.0033 0.00063)	-0.0014 (-0.0036 - 0.00072)	-0.0011 (-0.0033 - 0.0010)
	# Faculty at Program	0.00068 (0.00012 - 0.0012)*	0.00061 (0.00013 - 0.0012)*	0.00051 (0.00010 - 0.0011)
	>mean duty hours/day for PGY1	-0.17 (-0.34 - 0.0020)*	-0.17 (-0.35 - 0.0057)*	-0.17 (-0.36 - 0.0030)*
	% ambulatory care time for PGY1	-0.00097 (-0.0046 - 0.0027)	-0.00072 (-0.0043 - 0.0029)	-0.00096 (-.0046 - - 0.0027)
	# of teaching sites	0.0084 (-0.0063 - 0.023)	0.0049 (-0.0093 - 0.019)	0.0053 (-0.0088 - 0.019)
External Factors	Distance between program and SI (0.4 – 6 miles)		0.049 (-0.052 - 0.15)	0.032 (-0.068 -0.13)
	Distance between program and SI (6 – 37 miles)		0.084 (-0.048 - 0.21)	0.071 (-0.059 -0.20)
	Distance between program and SI (37 – 96 miles)		0.19 (-0.083 - 0.46)	0.17 (-0.092 - 0.43)
	Distance between program and SI (96+ miles)		-0.13 (-0.35 -0.081)	-0.16 (-0.38 - 0.064)
	SI in Academic Health Center		0.075 (-0.031 - 0.18)	0.035 (-0.069 - 0.14)
	Hospital beds at SI's primary teaching hospital (standardized)		-0.012 (-0.075 - 0.051)	-0.0085 (-0.073 - 0.056)
	Competing primary care programs at same SI		0.032 (-0.067 - 0.13)	0.054 (-0.045 - 0.15)
	Competing specialty care programs at same SI (standardized)		-0.0023 (-0.071- 0.067)	0.0085 (-0.060 - 0.077)
	Age of program (standardized)		0.0047 (-0.041 - 0.051)	-0.0019 (-0.050 - 0.046)
	SI Ownership Private/Nonprofit		0.0043 (-0.0077 - 0.086)	-0.0011 (-0.082 - 0.080)
	Total CME Medicare GME Funding Received (standardized)		0.0021 (-0.063 - 0.068)	-0.019 (-0.089 - 0.052)
	SI received Teaching Health Center funding		0.14 (-0.13 - 0.40)	0.17 (-0.081 - 0.43)
	Program with ACGME accreditation warnings		-0.21 (-0.46 - 0.032)	-0.21 (-0.47 - 0.040)

County Characteristic	% Medicare enrollees in SI's county			-1.27 (-2.61 - 0.026)*
	SI's county is non-metropolitan (rural)			0.13 (-0.12 - 0.39)
	Health Prof Shortage Area designation for all SI's county			-0.12 (-0.021 - -0.0039)**
	% under FPL in SI's county			0.0065 (-0.0029 - 0.015)
	% Uninsured in SI's county			-0.0094 (-0.019 - 0.00022)*
Constant		2.64 (2.47 - 2.81)	2.58 (2.39 -2.78)	2.89 (2.53 - 3.23)
$\sqrt{\psi}$		1.11 (1.03 - 1.19)	1.10 (1.02 - 1.19)	1.10 (1.02 - 1.18)
$\sqrt{\theta}$		0.39 (0.37 - 0.41)	0.39 (0.37 - 0.41)	0.39 (0.37 - 0.41)
AIC		7944.05	7952.92	7947.23
BIC		8027.98	8120.78	8147.37

*$p<.05$, **$p<.01$, ***$p<.001$

Figure 4.3: Unadjusted Clinical Competency Learning Rates

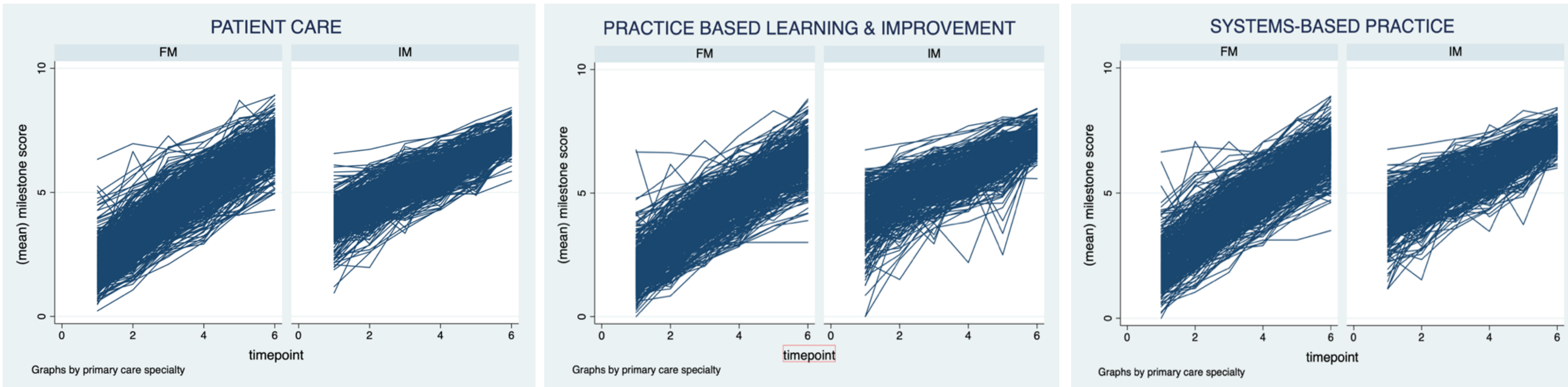
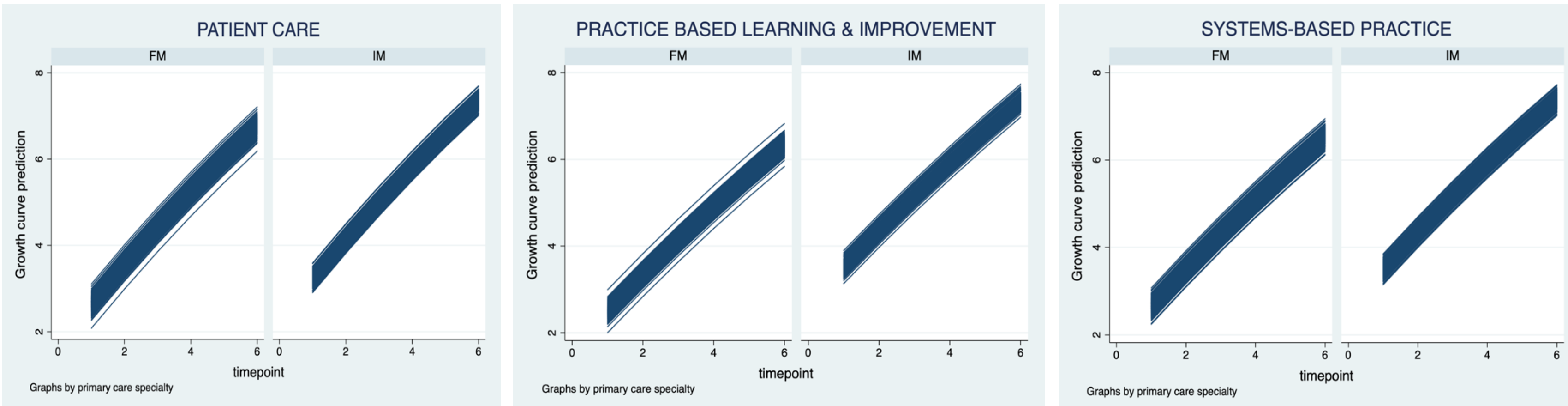


Figure 4.4: Predicted Clinical Competency Learning Rates



CHAPTER 5: AIM 3

The Reliability of Resident Performance Training in Family Medicine and Internal Medicine Accredited-Programs

Background

Competency-based medical education (CBME) is an outcomes-based training approach for physician professional development in Graduate Medical Education (GME).[3,40] In 2014, CBME assessments across all years of resident training were required by the Accreditation Council of GME (ACGME).[86] Yet current CBME assessments rarely incorporate practice- or systems-based measures and continue to rely on high-stakes licensing examinations to determine competence.[98,101]

For practicing physicians, audit and feedback of individual performance in health systems use publicly reported quality indicators to monitor and improve performance.[102,103] Empirical studies on audit and feedback argue about its potential to improve physician behavioral in the clinical microsystem.[104,105] Audit and feedback in GME could train residents to improve their individual clinical competency and collective performance in a health care system, a critical skill for future practice.[106] Yet these studies focus on the quality of care in Teaching Hospitals and inpatient-based settings.[107–111] For ambulatory care measures, the National Committee for Quality Assurance (NCQA)'s Healthcare Effectiveness Data and Information Set (HEDIS) is a common publicly available data source for outpatient-based care performance.[112] The potential to leverage publicly reported data could also assess the performance of residents and their associated GME programs.[47,92,113]

Publicly reported performance data are assumed to have adequate numbers of patients to enable the reliable measurement of individual physician performance, given the multi-level nature of this data.[114–116] Studies on group-level reliability of publicly available ambulatory care quality measures incorporate multi-payer and physician group performance initiatives, for which when pooled, produce sufficient patient samples.[117,118] However in GME, one of the most frequent accreditation warnings issued to family medicine programs is an insufficient number of patients cared for in ambulatory care settings.[73] Since no previous study has examined the reliability of quality indicators for individual GME resident and program performance assessment, understanding the number of patients needed to produce reliable quality scores during GME training is needed to incorporate audit and feedback for use in CBME GME.

This study explores the reliability of ambulatory care quality scores for family medicine and internal medicine residents and their GME programs by examining resident and program performance on HEDIS. Using data from 8 Family Medicine and Internal Medicine-accredited GME programs and sponsored by a large integrated health care system, the study examines the extent to which reliable estimates of resident performance are possible using available patient sample sizes from HEDIS scores.

Methods

Study Sample

The study sample includes resident performance for 31 HEDIS measures for 566 resident physicians who trained between 2014-2017 in one of 8 family medicine and internal medicine-accredited GME programs located in 6 medical centers operated by Kaiser Permanente Southern California (KPSCAL).

Data

KPSCAL resident quality indicators include publicly reported NCQA's HEDIS scores and routinely collected for KPSCAL patients. HEDIS data between June 2014-May 2017 from residents' panels of patients were extracted for analysis. HEDIS scores for each measure assigned to a resident are calculated based on the resident's patient encounter meeting the NCQA specification for the HEDIS measure over the total number of eligible patients. Calculation of residents' scores mirror performance standards used for practicing KPSCAL physicians. A census of residents' patients were obtained for 31 HEDIS measures for diabetes management, prevention health screening (cancer), cardiovascular health (blood pressure control) and monitoring of patients on persistent medications at the individual resident level. These measures were collapsed into 12 measures due to overlapping HEDIS categories. For the 12 HEDIS measures examined, most were under the NCQA domains for Effectiveness of Care and the most frequently reported NCQA HEDIS category were for Diabetes and Prevention & Screening.

To characterize the GME programs and patient population served, HEDIS data were linked at the GME program level to the 2016-17 American Medical Association (AMA) and American Association of Medical College's (AAMC)'s National GME Census, 2016 CMS Cost Reports. County level characteristics were linked to each training program's FIPS codes to the 2015-16 Area Health Resource File and the Robert Wood Johnson Foundation/University of Wisconsin Population Health Institute County Health Rankings, representing 5 counties for the 6 medical centers.

Methods: Statistical analysis

Frequency, mean, and spread (standard deviation) were calculated to describe program and county-level characteristics for residents and each program included in the study. One-way ANOVA tests were conducted to determined differences across program mean rates. Results with $p<0.05$ were considered statistically significant.

To decompose the patient-level variation attributable to differences between residents vs. within residents, we calculated intraclass-correlation (ICC) for each of the 12 HEDIS measures using the STATA command *loneway*.[119] ICC were calculated for each available reporting year (ie: 2014, 2015, 2016, 2017) using an unadjusted one-way, random-effects ANOVA model, which estimates the intra-unit correlation reliability of mean resident HEDIS rates based on available patient counts. Using the STATA command *sbrowni* for the Spearman-Brown Prophecy, we calculated patient sample sizes needed to achieve reliability alpha levels of 0.70 and 0.80 for each measurement year, which are widely accepted standards for determining unit-level reliability.[120,121] Differences were calculated between the required patient sample sizes for each reliability level for 0.70 and 0.80 physician-level reliability minus the available patient sample sizes for residents' HEDIS scores in the corresponding year. Results with $p<.05$ were considered statistically significant. Analyses were performed using STATA 15.1.[67] This study received a human subjects exemption from the Southern California Kaiser Permanente Institutional Review Board (IRB).

Results

Table 5.1 describes the sample of 8 family medicine and internal medicine GME programs in 6 medical centers according to program and county-level characteristics. The majority of programs are in family medicine (75%) with a mean and standard deviation (SD) age of 31.6 (15.9) years. Each program had a mean (SD) resident/faculty ratio of 0.7 (0.3). First year residents spent a mean (SD) of 24.5% (11.5) time in the ambulatory care setting. County-level characteristics where GME programs were located are comprised of 50% female, 43.5% Hispanic/Latino and bachelor's degrees for the highest degree earned (17.6%). 22.1% of the programs' county were uninsured, 13.5% were enrolled in Medicare, and 15.7% were living under the Federal Poverty Level. For estimated patient health-related characteristics, 11.2% were smokers, 22.5% were categorized as obese (body mass index > 30), and 8.5% were living with diabetes.

Table 5.2 describes the HEDIS score availability for residents, ranging from a count of 138 to 6,198. The majority of HEDIS scores for residents were for patients needing diabetic care and preventative screening tests. Of these, almost all residents had performed screening tests (>91.3% of residents). The mean (SD) number of patients seen/resident across all 12 HEDIS scores was 19 (13.6) patients over an mean (SD) of 365 (165.5) residents. Table 5.2 also compares resident HEDIS scores across each medical center. 9 of 12 measures (75%) were found to be highly statistically significant ($p<.001$), indicating notable differences of HEDIS performance across each medical center.

Table 5.3 describes the unadjusted ICC and sample size requirements for residents' HEDIS measures by each reporting year. With the exception of one HEDIS measure (Childhood Immunizations – Combination 3), ICC estimates were less than 0.06 due to resident-level variation. For Childhood Immunizations, the largest ICCs were observed for this measure in comparison to the other 11 HEDIS measures (0.35 to 0.44 over four reporting years). Reliability alpha estimates at 0.7 and the minimum number of required patients were met by approximately one-half of the 12 HEDIS measures, although the reporting year of 2014 only had 1 of 12 HEDIS measure that met the minimum sample size. For reliability estimates at the 0.8 alpha level, only colorectal screening tests met the minimum sample size for the reporting years of 2015, 2016, and 2017. For measures with the largest deficiencies in patient sample sizes, most were related to intermediate patient outcomes, including providing comprehensive diabetic care and controlling high blood pressure.

Table 5.4 reports the unadjusted ICC and sample size requirements for HEDIS measures aggregated at the medical center level by each reporting year. Similar to the resident-level data in Table 5.3, estimate ICC values were small (ICC < 0.03), with the exception of the Childhood Immunizations – Combination 3 measure (ICC = 0.18 to 0.30 over four years). Reliability alpha estimates at 0.7 for the minimum number of required patients were met by all the HEDIS measures, with the exception of controlling high blood pressure, which was only met in the 2014 reporting year. At the 0.8 alpha level, mostly all HEDIS measures met the minimum number of required patients needed, with the exception of the 2015 reporting year. Similar to the individual resident level, providing comprehensive diabetic care and controlling high blood pressure also had large insufficient patient sample sizes.

Discussion:

This study finds that annual measurement of individual resident performance scores on HEDIS measures is not reliable at the resident-level, with nearly all HEDIS measures requiring more patients than available in resident panels. Reliable measures found in resident performance were mostly for meeting preventive screening test recommendations. In contrast, resident performance for HEDIS intermediate outcome measures were found to have the largest deficiencies in patient sample sizes. When HEDIS measures were pooled across residents at the medical center level, however, higher estimates of reliability was achieved, with adequate patient sample sizes at the 0.7 alpha level indicating the potential to use HEDIS for group-level performance assessments. Thus, pooling residents' HEDIS scores by medical center produces more reliable measurement of GME program performance to the 12 HEDIS measures assessed. Medical center-level reporting of HEDIS performance is primarily used by health care systems; within an integrated health care system like Kaiser Permanente, medical center HEDIS performance assists with identifying practice-based gaps that could be improved via quality improvement programs.[122,123]

Caring for patients living with chronic conditions, including reducing hemoglobin A1c (HbA1c), cholesterol, and blood pressure levels, appear to challenge resident performance. These findings are to be generally expected since an individual resident's time in the ambulatory care setting to manage chronic illnesses can be limited, especially in the first year of residency. This concurs with previous reports on GME accreditation citations that describe family medicine GME programs' struggle scheduling resident continuity clinic time.[73] If HEDIS scores are incorporated for CBME GME, scheduling adequate resident time in the ambulatory care setting, selecting reliable HEDIS measures, and ensuring sufficient patient sample sizes for performance measurement presents the following challenges in GME programs: 1. Do programs allocate more time for residents in the ambulatory care setting in the interest of adequate patient sample sizes for individual resident performance assessment? 2. Do GME programs pool HEDIS scores at the program level and use these measures to monitor program performance? 3. Should programs only monitor and compare medical center and/or resident performance for HEDIS measures that have a minimum acceptable reliability of 0.70? These questions align with accrediting bodies and the National Academy of Medicine calling for better GME accountability and improved training to manage chronic conditions, and highlights the importance of ensuring reliable performance assessment of residents.[3,4,11]

The use of clinical quality scores for residents has been long argued as a potential source of data to use for resident assessment for clinical competency.[47,74,124–126] In tandem, the implementation of the ACGME Milestones Project adopted language to utilize quality improvement efforts and related quality scores to assess the core competencies in patient care, practice-based learning and improvement, and systems-based practice.[18] Whether HEDIS scores close resident-related practice gaps and whether they improve resident achievement of core competencies are future studies. Yet understanding the mechanisms of clinical quality scores for physicians in training, as this study has shown, is a fundamental step these measures for CBME GME assessment.

As the NCQA continues to utilize HEDIS compare health plans and systems across the nation, the utility of using HEDIS scores for GME is high, given its potential to assess resident and/or GME program level performance. A key example to improve resident level assessment is the movement to expand assessment methods beyond licensing examinations.[3,128–130] The National Board of Medical Examiners (NBME) is entertaining the elimination of quantitative scores for physician-in-training performance on licensing exams, thus reporting only a pass/fail score.[131] Thus, the role of quality scores such as HEDIS become more critical to understanding a more nuanced metric for resident performance, similar to audit and feedback. For assessing GME program level performance, understanding how GME programs who utilize quality scores could align GME program training with the same performance standards as practicing physicians. This is of particular interest to the National Academy of Medicine, who has called for a better understanding to identify outcome-based measures to align GME training with health systems performance.[3]

Limitations

This study's primary limitation is using performance data from a single integrated health care system and in a single region of California. Kaiser Permanente medical groups and its health plan consistently achieve high quality of care relative to other groups.[127] Our analytic data sample also did not incorporate patient-level characteristics to implement adjusted analytic models, which may also improve the precision of our results. These factors limit the generalizability of our findings to other health systems and practices that have different performance outcomes and patient case mixes.

Conclusion

Opportunities exist to utilize publicly reported quality data for GME programs. HEDIS measure reliability in ambulatory care for residents varies among ACGME-accredited primary care residency programs and requires further study to understand the role of audit and feedback for resident clinical competence, CBME GME, and GME program accountability.

Table 5.1: GME Program Characteristics by Medical Center

Program Characteristic	Center 1	Center 2	Center 3	Center 4	Center 5	Center 6	Overall
Family Medicine Program (n, %)	1 (50)	1 (50)	1 (100)	1 (100)	1 (100)	1 (100)	6 (75)
Age of Program in years	32	49	25	26	6	24	31.6 (15.9)
# of Residents/Program	24.4	35.5	24	18	18	18	25.1 (8.8)
# Faculty/Program	48.5	55.9	13	55	12	8	36.1 (34.1)
Resident/Faculty Ratio	0.5	0.6	1.8	0.3	1.5	2.3	0.7 (0.3)
% Outpatient time for first year residents (PGY1)	15	40	15	20	20	20	24.5 (11.5)
# Training experiences outside of primary teaching site	1	1.4	3	1	2	5	2.1 (1.3)
# Beds at primary Teaching Hospital	460	460	460	460	460	460	460
County of Medical Center Population Characteristics							
% Female	50	50.4	50.1	49.6	49.2	50.4	50%
% Race							
Hispanic/Latino (all)	49.2	47.7	33.7	45.5	32	47.7	43
Black	8.9	8.7	1.7	6.4	5.1	8.7	7
Other Race	21.6	21.8	14.5	20.5	13.6	21.8	19 (3)
Highest Degree Earned is Bachelor's Degree	10.8	18.7	23.2	12	21.5	18.7	18
% Uninsured	21.6	24.5	19.3	22.5	18.4	24.5	22
% Medicare enrollees	12.2	13.3	14.1	13.9	14.2	13.4	13
% Living under FPL	17.6	17.1	11.7	15.6	13.9	17.1	16
% Smokers	13.1	10.6	9.9	12.4	11.1	10.6	11
% Obese (BMI>30)	27.3	21.6	9.5	25.8	20.1	21.6	23
% Diagnosed with Diabetes	8.6	9	7.7	9	7.2	9	8

Table 5.2: GME Program HEDIS Scores by Measure, Resident, Patients, and Medical Center

HEDIS Score Measure	Measure Count (n)	Resident Count (n)	% of Residents	# of Patients eligible (mean, SD)	Center 1 Score (mean, SD)	Center 2 Score (mean, SD)	Center 3 Score (mean, SD)	Center 4 Score (mean, SD)	Center 5 Score (mean, SD)	Center 6 Score (mean, SD)	Overall Scores (mean, SD)
Annual Monitoring for Patients on Persistent Medication***	2,323	536	94.7%	29 (24)	88.4% (14.1)	85.2% (18.4)	83.8% (14.5)	82.3% (14.1)	87.5% (12.0)	84.7% (11.1)	85.4% (15.2)
Breast Cancer Screening***	2278	517	91.3%	22 (17)	79.0% (19.5)	82.8% (22.9)	78.9% (13.3)	81.3% (8.7)	78.7% (18.3)	79.7% (9.8)	80.4% (17.6)
Cervical Cancer Screening***	2414	531	93.8%	53 (53)	77.2% (21.4)	85.6% (17.9)	80.4% (9.7)	80.8% (7.0)	82.7% (12.9)	82.8% (6.1)	81.9% (15.3)
Childhood Immunizations - Combination 3***	543	167	29.5%	2 (2)	89.1% (24.7)	96.2% (16.0)	61.0% (45.2)	0.0% (.0)	62.1% (44.2)	90.6% (25.9)	79.8% (36.0)
Colorectal Cancer Screening***	6198	550	97.2%	32 (31)	68.0% (20.1)	78.7% (22.5)	69.9% (16.2)	72.4% (12.9)	77.1% (13.0)	77.2% (10.4)	74.3% (18.4)
Comprehensive Diabetes Care: HbA1c < 8.0	375	185	32.7%	13 (10)	59.9% (14.6)	59.2% (26.2)	53.3% (20.4)	60.1% (17.9)	65.8% (25.7)	62.9% (17.2)	59.4% (21.4)
Comprehensive Diabetes Care: HbA1c Testing***	375	185	32.7%	13 (10)	93.5% (12.4)	95.9% (8.2)	90.7% (9.5)	87.6% (16.8)	96.7% (7.0)	91.1% (9.4)	92.9% (11.1)
Comprehensive Diabetes Care: LDL-C Control	138	138	24.4%	13 (10)	62.5% (8.8)	60.1% (22.2)	53.3% (14.1)	57.6% (16.7)	53.7% (25.7)	63.1% (13.9)	59.0% (17.5)
Controlling High Blood Pressure***	2290	539	95.2%	14 (12)	79.5% (18.7)	83.7% (21.6)	82.3% (14.9)	81.2% (13.7)	84.6% (13.3)	81.7% (12.3)	82.3% (17.4)
HbA1c < 8.0%***	5233	343	60.6%	10 (10)	59.0% (24.7)	64.1% (32.7)	62.5% (24.8)	60.7% (24.5)	70.5% (25.6)	69.4% (20.9)	64.0% (27.1)
HbA1c < 9.0%***	5233	343	60.6%	10 (10)	72.0% (23.7)	76.3% (28.6)	75.4% (21.5)	74.3% (21.4)	81.1% (23.2)	81.5% (17.4)	76.4% (24.0)
HbA1c Testing	1876	343	60.6%	13 (13)	92.6% (13.3)	92.2% (16.4)	92.1% (11.6)	92.3% (10.3)	91.8% (18.5)	94.1% (6.9)	92.5% (13.9)

***p<.001

Table 5.3: One-way Analysis-of-variance Estimates for Intraclass-Correlation (unadjusted) and Sample Size Requirements for HEDIS Estimated for Resident-Level Reliability between 2014-2017

HEDIS Measure	2014				2015				2016				2017			
	2014 ICC	2014 mean	2014 α 0.7 (Difference)	2014 α 0.8 (Difference)	2015 ICC	2015 mean	2015 α 0.7 (Difference)	2015 α 0.8 (Difference)	2016 ICC	2016 mean	2016 α 0.7 (Difference)	2016 α 0.8 (Difference)	2017 ICC	2017 mean	2017 α 0.7 (Difference)	2017 α 0.8 (Difference)
Annual Monitoring for Patients on Persistent Medications	0.038	50	58 (-8)	100 (-50)	0.040	81	56 (25)	96 (-15)	0.033	63	68 (-5)	116 (53)	0.030	76	73 (3)	125 (-49)
Breast Cancer Screening	0.024	38	97 (-59)	167 (-129)	0.032	60	71 (-11)	121 (-61)	0.034	45	66 (-21)	113 (68)	0.025	59	89 (-30)	153 (-94)
Cervical Cancer Screening	0.018	93	125 (-32)	215 (-122)	0.019	143	120 (23)	205 (-62)	0.016	107	143 (-36)	246 (139)	0.019	146	119 (27)	204 (-58)
Childhood Immuniz. Combination 3	0.352	2	4 (-2)	7 (-5)	0.412	4	4 (0)	6 (-2)	0.438	4	3 (1)	6 (2)	0.346	4	4 (0)	7 (-3)
Colorectal Cancer Screening	0.029	93	78 (15)	134 (-41)	0.031	239	72 (167)	124 (115)	0.031	188	72 (116)	123 (-65)	0.022	249	100 (149)	172 (77)
Comp. Diabetes Care: HbA1c < 8.0	0.019	27	115 (-88)	198 (-171)	-	-	-	-	-	-	-	-	-	-	-	-
Comp. Diabetes Care: HbA1c Testing	0.016	27	141 (-114)	242 (-215)	-	-	-	-	-	-	-	-	-	-	-	-
Comp. Diabetes Care: LDL-C Control	0.000	13	18 (-5)	31 (-18)	-	-	-	-	-	-	-	-	-	-	-	-
Controlling High Blood Pressure	0.009	25	242 (-217)	416 (-391)	0.006	38	413 (-375)	708 (-670)	0.026	32	87 (-55)	150 (118)	0.013	37	175 (-138)	300 (-263)
HbA1c < 8.0%	0.047	26	48 (-22)	81 (-55)	0.051	69	43 (26)	73 (-4)	0.059	56	38 (18)	66 (10)	0.047	76	47 (29)	81 (-5)
HbA1c < 9.0%	0.052	26	42 (-16)	73 (-47)	0.052	69	42 (27)	72 (-3)	0.057	56	38 (18)	66 (10)	0.043	76	51 (25)	87 (-11)
HbA1c Testing	0.024	13	92 (-79	157 (-144)	0.021	35	108 (-73)	185 (-150)	0.020	28	114 (-86)	196 (168)	0.026	38	84 (-46)	145 (-107)

Table 5.4: One-way Analysis-of-variance Estimates for Intraclass-Correlation (unadjusted) and Sample Size Requirements for HEDIS Estimated for Medical Center-Level Reliability between 2014-2017

HEDIS Measure	2014				2015				2016				2017			
	2014 ICC	2014 mean	2014 α 0.7 (Difference)	2014 α 0.8 (Difference)	2015 ICC	2015 mean	2015 α 0.7 (Difference)	2015 α 0.8 (Difference)	2016 ICC	2016 mean	2016 α 0.7 (Difference)	2016 α 0.8 (Difference)	2017 ICC	2017 mean	2017 α 0.7 (Difference)	2017 α 0.8 (Difference)
Annual Monitoring for Patients on Persistent Medications	0.006	1952	394 (1158)	675 (1277)	0.016	1806	140 (1666)	241 (1565)	0.007	1412	317 (1095)	544 (868)	0.001	2025	1612 (413)	2763 (-738)
Breast Cancer Screening	0.005	1354	432 (922)	741 (613)	0.002	1319	888 (431)	1523 (-204)	0.004	1002	486 (516)	833 (169)	0.005	1560	409 (1151)	702 (858)
Cervical Cancer Screening	0.010	3543	214 (3329)	366 (3177)	0.007	3510	300 (3210)	514 (2996)	0.001	2578	1476 (1102)	2530 (48)	0.003	3858	732 (3126)	1254 (2604)
Childhood Immuniz. Combination 3	0.301	31	5 (26)	9 (22)	0.205	35	9 (26)	15 (20)	0.202	23	9 (14)	16 (7)	0.181	33	10 (23)	18 (15)
Colorectal Cancer Screening	0.011	3807	196 (3611)	337 (3470)	0.010	5735	224 (5511)	383 (5352)	0.013	4402	173 (4229)	296 (4106)	0.009	6716	235 (6481)	403 (6313)
Comp. Diabetes Care: HbA1c < 8.0	0.007	570	324 (246)	555 (15)	-	-	-	-	-	-	-	-	-	-	-	-
Comp. Diabetes Care: HbA1c Testing	0.009	570	232 (338)	397 (173)	-	-	-	-	-	-	-	-	-	-	-	-
Comp. Diabetes Care: LDL-C Control	0.002	234	803 (-569)	1376 (-1142)	-	-	-	-	-	-	-	-	-	-	-	-
Controlling High Blood Pressure	0.003	997	767 (230)	1316 (-319)	0.005	38	413 (-375)	708 (-670)	0.026	32	87 (-55)	150 (-188)	0.013	37	175 (138)	300 (-263)
HbA1c < 8.0%	0.023	407	96 (311)	164 (243)	0.016	1548	140 (1408)	241 (1307)	0.012	1236	183 (1053)	313 (923)	0.008	1990	265 (1725)	455 (1535)
HbA1c < 9.0%	0.031	407	72 (335)	123 (284)	0.017	1548	129 (1419)	221 (1327)	0.010	1236	231 (1005)	395 (841)	0.007	1990	290 (1700)	497 (1493)
HbA1c Testing	0.005	203	458 (-255)	786 (-583)	0.003	774	701 (73)	1202 (-428)	0.002	607	893 (-286)	1478 (-871)	0.001	995	2675 (-1680)	4586 (-3591)

CHAPTER 6: CONCLUSION

This dissertation is a collective empirical study on GME's competing organizational factors and performance, and helps policymakers understand the fragmented GME learning environment. By better describing the relationships among Family Medicine and Internal Medicine-accredited GME programs with their sponsoring institutions and external environment, the three empirical studies explored the association of organizational behavioral factors and performance within resident experiences to address the training challenges with our future primary care workforce, especially within a tenuous federal funding environment impacting the delivery of health care services for the U.S. population. As trainee experiences may have a downstream impact on patient care, the systematic study on primary care GME sponsoring institutions and their training programs informs policy on how GME could design optimal learning environments and improve the training of our future primary care workforce, especially in community-based and non-hospital settings, the most common delivery setting for health care services today.

REFERENCES

1. Chen C, Petterson S, Phillips R, Bazemore A, Mullan F. Spending patterns in region of residency training and subsequent expenditures for care provided by practicing physicians for Medicare beneficiaries. *Jama*. 2014;312(22):2385-2393. doi:10.1001/jama.2014.15973

2. Weinstein DF, Thibault GE. Illuminating Graduate Medical Education Outcomes in Order to Improve Them: *Acad Med*. 2018;93(7):975-978. doi:10.1097/ACM.0000000000002244

3. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division. *Graduate Medical Education Outcomes and Metrics: Proceedings of a Workshop*. (Martin P, Zindel M, Nass S, eds.). Washington, D.C.: National Academies Press; 2018. doi:10.17226/25003

4. National Academy of Medicine. *Graduate Medical Education That Meets the Nation's Health Needs*. Washington, D.C.: National Academies Press; 2014. doi:10.17226/18754

5. Ludmerer KM. *Let Me Heal: The Opportunity to Preserve Excellence in American Medicine*. Oxford University Press; 2015.

6. Chen C, Chen F, Mullan F. Teaching health centers: a new paradigm in graduate medical education. *Acad Med J Assoc Am Med Coll*. 2012;87(12):1752-1756. doi:10.1097/ACM.0b013e3182720f4d

7. Blanchard J, Petterson S, Bazemore A, Watkins K, Mullan F. Characteristics and Distribution of Graduate Medical Education Training Sites: Are We Missing Opportunities to Meet U.S. Health Workforce Needs? *Acad Med*. 2016;91(10):1416-1422. doi:10.1097/ACM.0000000000001184

8. Kesselheim JC, Cassel CK. Service: An Essential Component of Graduate Medical Education. *N Engl J Med*. 2013;368(6):500-501. doi:10.1056/NEJMp1214850

9. Walter D, Whitcomb ME. Venues for clinical education in internal medicine residency programs and their implications for future training. *Am J Med*. 1998;105(4):262-265.

10. Cleland J, Durning SJ. Education and service: how theories can help in understanding tensions. *Med Educ*. 0(0). doi:10.1111/medu.13738

11. Chen C, Petterson S, Phillips RL, Mullan F, Bazemore A, O'Donnell SD. Toward Graduate Medical Education (GME) Accountability: Measuring the Outcomes of GME Institutions. *Acad Med J Assoc Am Med Coll*. 2013;88(9):1267-1280. doi:10.1097/ACM.0b013e31829a3ce9

12. Turner TL, Fielder E, Ward MA. Balancing Service and Education in Residency Training: A Logical Fallacy. *JAMA Pediatr*. 2016;170(2):101. doi:10.1001/jamapediatrics.2015.3816

13. Holmboe ES. Competency-Based Medical Education and the Ghost of Kuhn: Reflections on the Messy and Meaningful Work of Transformation. *Acad Med J Assoc Am Med Coll*. 2017;(Journal Article). doi:10.1097/ACM.0000000000001866

14. Boex JR, Leahy PJ. Understanding residents' work: moving beyond counting hours to assessing educational value. *Acad Med J Assoc Am Med Coll*. 2003;78(9):939-944.

15. Brooks JV, Singer SJ, Rosenthal M, Chien AT, Peters AS. Feeling inadequate: Residents' stress and learning at primary care clinics in the United States. *Med Teach*. 2017;0(0):1-8. doi:10.1080/0142159X.2017.1413236

16. Sisson SD, Dalal D. Internal Medicine Residency Training on Topics in Ambulatory Care: A Status Report. *Am J Med*. 2011;124(1):86-90. doi:10.1016/j.amjmed.2010.09.007

17. Francis MD, Wieland ML, Drake S, et al. Clinic Design and Continuity in Internal Medicine Resident Clinics: Findings of the Educational Innovations Project Ambulatory Collaborative. *J Grad Med Educ*. 2015;7(1):36-41. doi:10.4300/JGME-D-14-00358.1

18. Holmboe ES, Edgar L, Hamstra SJ. *ACGME: The Milestones Guidebook*. 2016.

19. Edgar L, Hamstra S. ACGME Milestones Project: Lessons Learned and What's Next. 2016:43.

20. Arora V, Guardiano S, Donaldson D, Storch I, Hemstreet P. Closing the gap between internal medicine training and practice: Recommendations from recent graduates. *Am J Med*. 2005;118(6):680-685. doi:10.1016/j.amjmed.2005.03.022

21. Weinberger SE, Smith LG, Collier VU, for the Education Committee of the American College of Physicians*. Redesigning Training for Internal Medicine. *Ann Intern Med*. 2006;144(12):927. doi:10.7326/0003-4819-144-12-200606200-00124

22. Warm EJ, Schauer DP, Diers T, et al. The ambulatory long-block: an accreditation council for graduate medical education (ACGME) educational innovations project (EIP). *J Gen Intern Med*. 2008;23(7):921-926. doi:10.1007/s11606-008-0588-y

23. Bowen JL, Hirsh D, Aagaard E, et al. Advancing educational continuity in primary care residencies: an opportunity for patient-centered medical homes. *Acad Med J Assoc Am Med Coll*. 2015;90(5):587-593. doi:10.1097/ACM.0000000000000589

24. Weinstein DF, Co JPT. Predictions and Prescriptions for Institutions Sponsoring Graduate Medical Education. *J Grad Med Educ*. 2017;9(6):680-682. doi:10.4300/JGME-D-17-00729.1

25. Thornton PH, Ocasio William, Lounsbury Michael. The institutional logics perspective : foundations, research, and theoretical elaboration. http://public.eblib.com/choice/publicfullrecord.aspx?p=845919. Published 2012.

26. Haveman HA, Gualtieri G. *Institutional Logics*. Vol 1. Oxford University Press; 2017. doi:10.1093/acrefore/9780190224851.013.137

27. Friedland R, Alford RR. Bringing society back in: Symbols, practices and institutional contradictions. 1991.

28. Dunn MB, Jones C. Institutional Logics and Institutional Pluralism: The Contestation of Care and Science Logics in Medical Education, 1967–2005. *Adm Sci Q*. 2010;55(1):114-149. doi:10.2189/asqu.2010.55.1.114

29. Daft RL, Murphy J, Wilmott H. *Organization Theory & Design : An International Perspective*.; 2017.

30. Shortell SM, Kaluzny AD. *Health Care Management: Organization, Design, and Behavior*. Thomson Delmar Learning; 2006.

31. Gehman J. How Institutions Matter: From the Micro Foundations of Institutional Impacts to the Macro Consequences of Institutional Arrangements. *Res Sociol Organ*. 2016;48:1-34.

32. Yeager VA, Menachemi N, Savage GT, Ginter PM, Sen BP, Beitsch LM. Using resource dependency theory to measure the environment in health care organizational studies: a systematic review of the literature. *Health Care Manage Rev*. 2014;39(1):50-65. doi:10.1097/HMR.0b013e3182826624

33. Pfeffer J, Salancik G. *The External Control of Organizations*. Harper & Row; 1978.

34. Lawrence PR, Lorsch JW. Differentiation and Integration in Complex Organizations. *Adm Sci Q*. 1967;12(1).

35. Levitt B, March JG. *Organizational Learning*. In: March JG, eds. *The Pursuit of Organizational Intelligence*. 1st ed. Malden, MA: Blackwell Publishing Inc. 1999.

36. Scott WR, Davis GF. *Organizations and Organizing: Rational, Natural and Open System Perspectives*. New York, NY: Routledge/Taylor & Francis Group; 2007.

37. Scott WR, Ruef M, Mendel PJ, Caronna CA. *Institutional Change and Healthcare Organizations: From Professional Dominance to Managed Care*. University of Chicago Press; 2000. https://books.google.com/books?id=vomaHpFVcOAC.

38. Pisano G, Boomer R, Edmonson A. Organizational Differences in Rates of Learning: Evidence from the Adoption of Minimally Invasive Cardiac Surgery. *Manag Sci*. 2001;47(6):752-768. doi:10.1287/mnsc.47.6.752.9811

39. Pusic MV, Boutis K, Hatala R, Cook DA. Learning curves in health professions education. *Acad Med J Assoc Am Med Coll*. 2015;90(8):1034-1042. doi:10.1097/ACM.0000000000000681

40. Holmboe ES. Realizing the promise of competency-based medical education. *Acad Med J Assoc Am Med Coll*. 2015;90(4):411-413. doi:10.1097/ACM.0000000000000515

41. Rui P, Okeyode T. National Ambulatory Medical Care Survey: 2015 State and National Summary Tables. 2015:35.

42. ACGME. *ACGME Program Requirements for Graduate Medical Education in Family Medicine*.; 2016.

43. ACGME. *ACGME Program Requirements for Graduate Medical Education in Internal Medicine*.; 2008.

44. Durfey SNM, George P, Adashi EY. Permanent GME Funding for Teaching Health Centers. *JAMA*. 2017;317(22):2277. doi:10.1001/jama.2017.5298

45. Gupta R, Dubé K, Bodenheimer T. The Road to Excellence for Primary Care Resident Teaching Clinics. *Acad Med J Assoc Am Med Coll*. 2016;91(4):458-461. doi:10.1097/ACM.0000000000001100

46. Bodenheimer T, Ghorob A, Willard-Grace R, Grumbach K. The 10 building blocks of high-performing primary care. *Ann Fam Med*. 2014;12(2):166-171. doi:10.1370/afm.1616

47. Holmboe ES, Prince L, Green M. Teaching and improving quality of care in a primary care internal medicine residency clinic. *Acad Med J Assoc Am Med Coll*. 2005;80(6):571-577.

48. Morris CG, Lesko SE, Andrilla HA, Chen FM. Family Medicine Residency Training in Community Health Centers: A National Survey: *Acad Med*. 2010;85(10):1640-1644. doi:10.1097/ACM.0b013e3181f08e2b

49. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. *Milbank Q*. 2005;83(3):457-502.

50. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. *Milbank Q*. 1996;74(4):511-544.

51. Hirsh DA, Ogur B, Thibault GE, Cox M. "Continuity" as an organizing principle for clinical education reform. *N Engl J Med*. 2007;356(8):858-866.

52. Johnson RK, Swain M, eds. *Immersion Education: International Perspectives*. Cambridge: Cambridge University Press; 1997. doi:10.1017/CBO9781139524667

53. Dulay M, Laponis R, O'Brien BC, Gupta R, Ramanan RA, Julian K. An Intensive Continuity Clinic Immersion Experience for Interns: A Springboard to Confidence and Satisfaction With Continuity Clinic. *J Grad Med Educ*. 2017;9(5):622-626. doi:10.4300/JGME-D-16-00696.1

54. Cruess RL, Cruess SR, Boudreau JD, Snell L, Steinert Y. A Schematic Representation of the Professional Identity Formation and Socialization of Medical Students and Residents: A

Guide for Medical Educators. *Acad Med J Assoc Am Med Coll*. 2015;(Journal Article). doi:10.1097/ACM.0000000000000700

55. Kassirer JP. Redesigning Graduate Medical Education — Location and Content. *N Engl J Med*. 1996;335(7):507-509. doi:10.1056/NEJM199608153350709

56. Stepczynski J, Holt SR, Ellman MS, Tobin D, Doolittle BR. Factors Affecting Resident Satisfaction in Continuity Clinic-a Systematic Review. *J Gen Intern Med*. 2018;33(8):1386-1393. doi:10.1007/s11606-018-4469-8

57. Asch DA, Nicholson S, Srinivas S, Herrin J, Epstein AJ. Evaluating obstetrical residency programs using patient outcomes. *Jama*. 2009;302(12):1277-1283. doi:10.1001/jama.2009.1356

58. Medicare C for, Baltimore MS 7500 SB, Usa M. Cost Reports by Fiscal Year. https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Cost-Reports/Cost-Reports-by-Fiscal-Year.html. Published November 12, 2014. Accessed November 8, 2018.

59. HRSA. HRSA Data Warehouse – Area Health Resource Files. 2017. https://datawarehouse.hrsa.gov/topics/ahrf.aspx.

60. American Association of Teaching Health Centers. http://aathc.org. Accessed March 7, 2019.

61. Brotherton SE, Etzel SI. Graduate Medical Education, 2016-2017. *JAMA*. 2017;318(23):2368-2387.

62. American College of Physicians. Internal Medicine vs. Family Medicine | ACP. https://www.acponline.org/about-acp/about-internal-medicine/career-paths/medical-student-career-path/internal-medicine-vs-family-medicine. Accessed February 12, 2019.

63. American Medical Association. FREIDA Residency Program Database | Medical Fellowship Database | AMA. https://freida.ama-assn.org/Freida/#/. Accessed February 12, 2019.

64. Weber S, Péclat M. Stata Journal | Article. *Simple Command Calc Travel Distance Travel Time*. 17(4). https://www.stata-journal.com/article.html?article=dm0092. Accessed November 8, 2018.

65. Rabe-Hesketh S, Skrondal A. *Multilevel and Longitudinal Modeling Using Stata (Third Edition). Volume I: Continuous Responses*. College Station, TX: Stata Press; 2012.

66. Rabe-Hesketh S, Everitt B. *A Handbook of Statistical Analyses Using Stata*. 4th ed. CRC Press; 2006. https://doi.org/10.1111/j.1541-0420.2006.00787_10.x. Accessed March 3, 2019.

67. StataCorp. *Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC*.; 2017.

68. ArcGIS. ArcGIS Online | Interactive Maps Connecting People, Locations & Data. https://www.arcgis.com/index.html. Accessed February 12, 2019.

69. Johnston SC. The US Training System for Physicians—Need for Deeper Analysis. *JAMA*. 2018;320(10):982-983. doi:10.1001/jama.2018.12879

70. Barnes K, Morris CG. Clinic First: Prioritizing Primary Care Outpatient Training for Family Medicine Residents at Group Health Cooperative. *J Gen Intern Med*. 2015;30(10):1557-1560. doi:10.1007/s11606-015-3272-z

71. How Primary Care Affects Health Care Costs and Outcomes | The U.S. Senate Committee on Health, Education, Labor & Pensions. https://www.help.senate.gov/hearings/how-primary-care-affects-health-care-costs-and-outcomes. Accessed February 9, 2019.

72. Detert JR, Edmondson AC. No Exit, No Voice: The Bind of Risky Voice Opportunities in Organizations. *Acad Manag Proc*. 2005;2005(1):O1-O6. doi:10.5465/ambpp.2005.18780787

73. Pugno PA, Epperly TD. Residency Review Committee for Family Medicine: an analysis of program citations. *Fam Med*. 2005;37(3):174-177.

74. Smirnova A, Sebok-Syer SS, Chahine S, et al. Defining and Adopting Clinical Performance Measures in Graduate Medical Education: Where Are We Now and Where Are We Going? *Acad Med*. January 2019:1. doi:10.1097/ACM.0000000000002620

75. Wynn B. Opening the "Black Box" of GME Costs and Benefits: A Conceptual Model and a Call for Systematic Studies. *J Grad Med Educ*. 2015;7(1):125-127. doi:10.4300/JGME-D-14-00751.1

76. Rittenhouse DR, Ament A, Grumbach K, Petterson S, Levin Z, Bazemore A. Guide to Graduate Medical Education Funding in California. *Calif Health Care Found*. September 2018:27.

77. Bodenheimer TS, Smith MD. Primary Care: Proposed Solutions To The Physician Shortage Without Training More Physicians. *Health Aff (Millwood)*. 2013;32(11):1881-1886. doi:10.1377/hlthaff.2013.0234

78. Lucey CR. Medical Education: Part of the Problem and Part of the Solution. *JAMA Intern Med*. 2013;173(17):1639-1643. doi:10.1001/jamainternmed.2013.9074

79. Irby DM, Cooke M, O'Brien BC. Calls for reform of medical education by the Carnegie Foundation for the Advancement of Teaching: 1910 and 2010. *Acad Med J Assoc Am Med Coll*. 2010;85(2):220-227. doi:10.1097/ACM.0b013e3181c88449

80. Institute of Medicine (U.S.), ed. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, D.C: National Academy Press; 2001.

81. Holmboe E. The ACGME Milestones Guidebook: Competency-Based Training. 2016:41.

82. Merton R. King. *Social Research and the Practicing Professions*. Cambridge, Ma: ABT Books.; 1982.

83. Asch DA, Nicholson S, Srinivas SK, Herrin J, Epstein AJ. How do you deliver a good obstetrician? Outcome-based evaluation of medical education. *Acad Med J Assoc Am Med Coll*. 2014;89(1):24-26. doi:10.1097/ACM.0000000000000067

84. Selby JV, Schmittdiel JA, Lee J, et al. Meaningful variation in performance: what does variation in quality tell us about improving quality? *Med Care*. 2010;48(2):133-139. doi:10.1097/MLR.0b013e3181c15a6e

85. Bazzoli GJ, Shortell SM, Dubbs N, Chan C, Kralovec P. A taxonomy of health networks and systems: bringing order out of chaos. *Health Serv Res*. 1999;33(6):1683-1717.

86. Hamstra SJ, Edgar L, Yamazaki K, Holmboe ES. Milestones Annual Report 2017. 2017:387.

87. Curran PJ, Obeidat K, Losardo D. Twelve Frequently Asked Questions About Growth Curve Modeling. *J Cogn Dev*. 2010;11(2):121-136. doi:10.1080/15248371003699969

88. Kim CS, Lukela MP, Parekh VI, et al. Teaching Internal Medicine Residents Quality Improvement and Patient Safety: A Lean Thinking Approach. *Am J Med Qual*. 2010;25(3):211-217. doi:10.1177/1062860609357466

89. AHA. Regulatory Overload Report. Assessing the Regulatory Burden on Health Systems, Hospitals and Post-acute Care Providers. American Hospital Association. https://www.aha.org/guidesreports/2017-11-03-regulatory-overload-report. Published October 2017. Accessed June 16, 2019.

90. Welch WP, Bindman AB. Town and Gown Differences Among the 100 Largest Medical Groups in the United States: *Acad Med*. 2016;91(7):1007-1014. doi:10.1097/ACM.0000000000001240

91. Colbert CY, Ogden PE, Ownby AR, Bowe C. Systems-based practice in graduate medical education: systems thinking as the missing foundational construct. *Teach Learn Med*. 2011;23(2):179-185. doi:10.1080/10401334.2011.561758

92. Johnson JK, Miller SH, Horowitz SD. Systems-Based Practice: Improving the Safety and Quality of Patient Care by Recognizing and Improving the Systems in Which We Work. In: Henriksen K, Battles JB, Keyes MA, Grady ML, eds. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 2: Culture and Redesign)*. Advances in Patient Safety. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008. http://www.ncbi.nlm.nih.gov/books/NBK43731/. Accessed August 16, 2018.

93. Duffy FD, Holmboe ES. Competence in Improving Systems of Care Through Practice-Based Learning and Improvement. In: *Practical Guide to the Evaluation of Clinical Competence*. 1st ed. Philadelphia: Elsevier; 2008.

94. Frenk J, Chen L, Bhutta ZA, et al. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. *The Lancet*. 2010;376(9756):1923-1958. doi:10.1016/S0140-6736(10)61854-5

95. Asch DA, Bilimoria KY, Desai SV. Resident Duty Hours and Medical Education Policy - Raising the Evidence Bar. *N Engl J Med*. 2017;376(18):1704-1706. doi:10.1056/NEJMp1703690

96. Philibert I, Friedmann P, Williams WT, ACGME Work Group on Resident Duty Hours. Accreditation Council for Graduate Medical Education. New requirements for resident duty hours. *Jama*. 2002;288(9):1112-1114.

97. Sen S, Kranzler HR, Didwania AK, et al. Effects of the 2011 Duty Hour Reforms on Interns and Their Patients: A Prospective Longitudinal Cohort Study. *JAMA Intern Med*. 2013;173(8):657-662. doi:10.1001/jamainternmed.2013.351

98. Hauer KE, Vandergrift J, Hess B, et al. Correlations Between Ratings on the Resident Annual Evaluation Summary and the Internal Medicine Milestones and Association With ABIM Certification Examination Scores Among US Internal Medicine Residents, 2013-2014. *JAMA*. 2016;316(21):2253. doi:10.1001/jama.2016.17357

99. Peabody MR, O'Neill TR, Peterson LE. Examining the Functioning and Reliability of the Family Medicine Milestones. *J Grad Med Educ*. 2017;9(1):46-53. doi:10.4300/JGME-D-16-00172.1

100. Hauer KE, Vandergrift J, Lipner RS, Holmboe ES, Hood S, McDonald FS. National Internal Medicine Milestone Ratings: Validity Evidence From Longitudinal Three-Year Follow-Up. *Acad Med*. 2018;Publish Ahead of Print. doi:10.1097/ACM.0000000000002234

101. Hoffman KI. The USMLE, the NBME subject examinations, and assessment of individual academic achievement. *Acad Med J Assoc Am Med Coll*. 1993;68(10):740-747.

102. Colquhoun H, Michie S, Sales A, et al. Reporting and design elements of audit and feedback interventions: a secondary review: Table 1. *BMJ Qual Saf*. 2017;26(1):54-60. doi:10.1136/bmjqs-2015-005004

103. Gude WT, Brown B, van der Veer SN, et al. Clinical performance comparators in audit and feedback: a review of theory and evidence. *Implement Sci*. 2019;14(1):39. doi:10.1186/s13012-019-0887-1

104. Gude WT, van Engen-Verheul MM, van der Veer SN, et al. Effect of a web-based audit and feedback intervention with outreach visits on the clinical performance of multidisciplinary

teams: a cluster-randomized trial in cardiac rehabilitation. *Implement Sci*. 2016;11(1). doi:10.1186/s13012-016-0516-1

105. Hysong SJ, Kell HJ, Petersen LA, Campbell BA, Trautner BW. Theory-based and evidence-based design of audit and feedback programmes: examples from two clinical intervention studies. *BMJ Qual Saf*. 2017;26(4):323-334. doi:10.1136/bmjqs-2015-004796

106. Hwang AS, Harding AS, Chang Y, O'Keefe SM, Horn DM, Clark AL. An Audit and Feedback Intervention to Improve Internal Medicine Residents' Performance on Ambulatory Quality Measures: A Randomized Controlled Trial. *Popul Health Manag*. April 2019. doi:10.1089/pop.2018.0217

107. Lynn L, Hess BJ, Weng W, Lipner RS, Holmboe ES. Gaps in quality of diabetes care in internal medicine residency clinics suggest the need for better ambulatory care training. *Health Aff Proj Hope*. 2012;31(1):150-158. doi:10.1377/hlthaff.2011.0907

108. Burke LG, Frakt AB, Khullar D, Orav EJ, Jha AK. Association Between Teaching Status and Mortality in US Hospitals. *JAMA*. 2017;317(20):2105-2113. doi:10.1001/jama.2017.5702

109. Perez JAJ, Awar M, Nezamabadi A, et al. Comparison of Direct Patient Care Costs and Quality Outcomes of the Teaching and Non-Teaching Hospitalist Services at a Large Academic Medical Center. *Acad Med. 2018;93(3):491-497*. doi:10.1097/ACM.0000000000002026

110. Taylor DH, Whellan DJ, Sloan FA. Effects of Admission to a Teaching Hospital on the Cost and Quality of Care for Medicare Beneficiaries. *N Engl J Med*. 1999;340(4):293-299. doi:10.1056/NEJM199901283400408

111. Ayanian JZ, Weissman JS. Teaching Hospitals and Quality of Care: A Review of the Literature. *Milbank Q*. 2002;80(3):569-593. doi:10.1111/1468-0009.00023

112. NCQA. HEDIS Measures and Technical Resources. NCQA. https://www.ncqa.org/hedis/measures/. Accessed June 24, 2019.

113. Edwards ST, Kim H, Shull S, Hooker ER, Niederhausen M, Tuepker A. Quality of Outpatient Care With Internal Medicine Residents vs Attending Physicians in Veterans Affairs Primary Care Clinics. *JAMA Intern Med*. 2019;179(5):711-713. doi:10.1001/jamainternmed.2018.8624

114. Scholle SH, Roski J, Dunn DL, et al. Availability of data for measuring physician quality performance. *Am J Manag Care*. 2009;15(1):67-72.

115. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The Unreliability of Individual Physician Report Cards for Assessing the Costs and Quality of Care of a Chronic Disease. *JAMA*. 1999;281(22):2098-2105. doi:10.1001/jama.281.22.2098

116. Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL. Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. *Ann Intern Med*. 2002;136(2):111-121. doi:10.7326/0003-4819-136-2-200201150-00008

117. Safran DG, Karp M, Coltin K, et al. Measuring patients' experiences with individual primary care physicians. Results of a statewide demonstration project. *J Gen Intern Med*. 2006;21(1):13-21. doi:10.1111/j.1525-1497.2005.00311.x

118. Sequist TD, Schneider EC, Li A, Rogers WH, Safran DG. Reliability of medical group and physician performance measurement in the primary care setting. *Med Care*. 2011;49(2):126-131. doi:10.1097/MLR.0b013e3181d5690f

119. StataCorp. *Loneway — Large One-Way ANOVA, Random Effects, and Reliability*.

120. Caci HM. *SBROWNI: Stata Module to Calculate Spearman-Brown Reliability Correction for Test Length*. Boston College Department of Economics; 1998. https://ideas.repec.org/c/boc/bocode/s351002.html. Accessed June 24, 2019.

121. Nunnally JC, Bernstein IH. *Psychometric Theory*. New York: McGraw-Hill; 1994.

122. Hu J, Schreiber M, Jordan J, George DL, Nerenz D. Associations Between Community Sociodemographics and Performance in HEDIS Quality Measures: A Study of 22 Medical Centers in a Primary Care Network. *Am J Med Qual Off J Am Coll Med Qual*. 2018;33(1):5-13. doi:10.1177/1062860617695456

123. Kanter MH, Lindsay G, Bellows J, Chase A. Complete Care at Kaiser Permanente: Transforming Chronic and Preventive Care. *Jt Comm J Qual Patient Saf*. 2013;39(11):484-494. doi:10.1016/S1553-7250(13)39064-3

124. ACGME Milestones, Evaluation & Research Department. ACGME Experience with Quality Scores. June 2018.

125. Schumacher DJ, Holmboe ES, van der Vleuten C, Busari JO, Carraccio C. Developing Resident-Sensitive Quality Measures: A Model From Pediatric Emergency Medicine. *Acad Med*. 2017;Publish Ahead of Print. doi:10.1097/ACM.0000000000002093

126. Mladenovic J, Shea JA, Duffy FD, Lynn LA, Holmboe ES, Lipner RS. Variation in Internal Medicine Residency Clinic Practices: Assessing Practice Environments and Quality of Care. *J Gen Intern Med*. 2008;23(7):914-920. doi:10.1007/s11606-008-0511-6

127. Care Beyond Compare: Kaiser Permanente Top-Rated by NCQA. *Perm Med*. September 2018. https://permanente.org/care-beyond-compare-kaiser-permanente-top-rated-ncqa/. Accessed November 10, 2018.

128. Berner ES, Brooks CM, Erdmann JB. Use of the USMLE to select residents. *Acad Med J Assoc Am Med Coll*. 1993;68(10):753-759.

129. Paniagua M, Salt J, Swygert K, Barone MA. Perceived Utility of the USMLE Step 2 Clinical Skills Examination from a GME Perspective. *J Med Regul*. 2018;104(2):51-57. doi:10.30770/2572-1852-104.2.51

130. Haist SA, Katsufrakis PJ, Dillon GF. The Evolution of the United States Medical Licensing Examination (USMLE): Enhancing Assessment of Practice-Related Competencies. *JAMA*. 2013;310(21):2245. doi:10.1001/jama.2013.282328

131. United States Medical Licensing Examination | Invitational Conference on USMLE Scoring. https://www.usmle.org/usmlescoring/. Accessed June 24, 2019.