Association Between Teaching Clinic Structure and the Readiness of Ophthalmology Residents to Enter Independent Practice

T. Michael Kashner, PhD, JD, MPH, [*,†] *Marjorie A. Bowman, MD, MPA,* [*,‡,§] *Catherine P. Kaminetzky, MD, MPH,* [*,‖] *Andrea D. Birnbaum, MD, PhD,* [*,¶] *John M. Byrne, DO,* [*,†] *Paul B. Greenberg, MD, MPH,* [**,††] *Steven S. Henley, MS, BS, BA,* [†,‡‡] *and Karen M. Sanders, MD* [*,§§]

[*]Department of Veterans Affairs, Office of Academic Affiliations, Washington, DC; [†]Department of Medicine, Loma Linda University School of Medicine, Loma Linda, CA; [‡]Family Medicine and Community Health, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA; [§]Family Medicine, Wright State University Boonshoft School of Medicine, Dayton, OH; [‖]Department of Medicine, University of Washington School of Medicine, Seattle, WA; [¶]Department of Ophthalmology, Northwestern University Feinberg School of Medicine, Chicago, IL; [**]Surgery Service, VA Providence Healthcare System, Providence, RI; [††]Department of Surgery (Ophthalmology), the Warren Alpert Medical School of Brown University, Providence RI; [‡‡]Martingale Research Corporation, Plano, TX; and [§§]Department of Internal Medicine, Virginia Commonwealth University School of Medicine, Richmond, VA

OBJECTIVE: Our objective is to determine if the structure of Graduate Medical Education teaching clinics is associated with how well ophthalmology residents are prepared to meet the workload demands of independent clinical practice.

DESIGN: Resident preparedness to enter independent practice was measured by the Readiness Index. Part of the Department of Veterans Affairs' new Workload-based Resident Academic Performance measures (WRAP), resident readiness is computed from electronic health records for residents by clinic and service-date. The index compares resident productivity net of supervision and adjusted for care quality to the average productivity of non-supervising ophthalmologists. Readiness comprises a Workload component (ratio of resident gross productivity to the average productivity of non-supervising ophthalmologists) and Supervision component (ratio of resident net of supervision to gross productivity). Teaching clinic factors include resident postgraduate-year level, resident-to-physician staff ratios, patient care complexity, and program size. Covariates include time into the academic year, facility quality ranking and complexity rating, and attending physician productivity rate.

SETTING: Study setting is 109 ophthalmology outpatient clinics from the United States Department of Veterans Affairs and its 1,300 annual ophthalmology resident positions rotating on 84,600 ophthalmology clinic-days during academic years from July 1, 2015, through June 30, 2019.

PARTICIPANTS: An average 2.6 residents at a second-year or higher saw 25.0 patients requiring 93.6 relative value units (RVUs) of workload.

RESULTS: Senior ophthalmology residents from clinics with higher resident-to-physician ratios had greater practice readiness than their counterparts primarily from having greater progressive autonomy from supervision. Residents from larger programs treating more complex patients had only slightly greater practice readiness than their counterparts primarily from having greater workload productivity.

CONCLUSIONS: The readiness of ophthalmology residents to enter independent practice is associated with their academic level and resident-to-physician staff ratios, and to a lesser extent care complexity and

Correspondence: Inquiries to T. Michael Kashner, PhD, JD, MPH, Department of Veterans Affairs, Office of Academic Affiliations, 810 Vermont Ave., NW, Washington, DC, 20420; Department of Medicine, Loma Linda University School of Medicine, 111 E. Sunset Dr. South, Redlands, CA 92373; e-mail: michael.kashner@va.go

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
https://doi.org/10.1016/j.jsurg.2024.08.020

program size. (J Surg Ed 81:103270. Published by Elsevier Inc. on behalf of Association of Program Directors in Surgery. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/))

KEY WORDS: Graduate medical education, resident readiness to enter independent practice, resident progressive autonomy from supervision, resident clinical workload productivity

INTRODUCTION

The supervision of resident physicians is the cornerstone of graduate medical education (GME) and is essential to advancing residents' clinical skills and medical knowledge while protecting patient safety and care effectiveness.[1] Recent studies, however, have echoed the academic community's concern that surgery residents are not being given the progressive autonomy from supervision they need to prepare them for independent practice.[2-4] Such unreadiness will exacerbate projected shortages of qualified surgeons in the U.S. currently estimated to be between 15,800 and 30,200 by 2034.[5] However, there is a dearth of studies examining whether residents' academic level, program size, professional staffing, and patient complexity of GME teaching clinics may enhance, or hinder, workload readiness of residents to enter independent practice.

In this paper, we investigated how these teaching ophthalmology clinic factors are associated with the readiness of ophthalmology residents to enter independent practice. Outpatient ophthalmology was selected because of the severe anticipated shortages of ophthalmologists in the coming decade,[6] the importance of outpatient care in GME training,[7] and the scientific need to reduce the variance across surgery specialties in workload production,[7] and outcomes of resident supervision.[4] These GME factors were selected because they can be objectively measured; are alterable by academic leaders; reflect the demands that supervising responsibilities place on attending physicians; and captures information about the resident, program, patient, and attending physician.

We define clinic "workload" in terms of the complexity and volume of medical procedures and clinic visits, quantified as billable charges, and computed in relative value units (RVU). We define "workload readiness" as the degree a resident's workload productivity will meet the clinical production demands of independent practice. We introduce the Department of Veterans Affairs (VA) new Workload-based Resident Academic Performance (WRAP) measures where workload readiness is computed on continuous scales from patient electronic health records. Readiness consists of 2 components: resident progressive autonomy from supervision and resident clinical workload productivity.

Our working assumptions come from the progressive independence hypothesis.[8-10] First, senior residents and those caring for less complex patient cases may require less supervision than their counterparts. Second, attending ophthalmologists who shift workload production to their residents and who have fewer residents to supervise may spend more time supervising each resident.

METHODS

Study Setting and Sample

The study setting is ophthalmology clinics at VA teaching medical centers. VA funds over 45,000 resident positions making it the second largest funder of GME behind the Centers for Medicare and Medicaid Services.[11] During the 2019 to 2020 academic year, VA medical centers hosted 118,012 health professions trainees including 47,521 physician residents and fellows.[12] There were 1,199 ophthalmology residents representing 11.9% of VA's 10,071 surgery residents.

VA medical centers are under consistent management and uniform policies governing resident supervision, where their success in balancing surgery resident progressive autonomy and patient outcomes was recently demonstrated.[4] We assume that: (1) attending physicians are responsible for assigning workload to residents whether by deliberate action or clinic process; (2) attending physicians may see other clinic patients while supervising residents; and (3) entrusting residents with progressive autonomy from supervision does not mean residents were unsupervised.

The study sample included all clinic-days at VA ophthalmology teaching clinics during 4 academic years from July 1, 2015, through June 30, 2019. "Clinic-days" are defined as having an attending ophthalmologist present for at least a half-day shift who provides at least 1 patient encounter with a nonzero RVU. A clinic-day is "teaching" whenever a PGY02 or higher resident is listed in the health record as a care provider on at least 1 patient encounter. To ensure sufficient statistical power to detect differences across clinics, a "teaching clinic" for a given academic year is eligible only if 10% or a minimum of 52 of its clinic-days were teaching clinic-days, whichever is larger.

Conceptual Framework: Outcomes

Study outcomes are resident readiness to enter independent practice and its components: resident produced workload and progressive autonomy from supervision. These indices are computed by clinic and service date from the metrics and rates comprising VA's Workload-based Resident Academic Performance measures.

Metrics:

(M1) *Resident Clinic-Workload* (**RCW**) equals total clinic RVUs per resident-day and represents the pool of available patient encounters that potentially can make up the resident's clinical experience.
(M2) *Resident Assigned-Workload* (**RAW**) equals clinic RVUs that residents were assigned per resident-day. Assigned-RVUs are those that a resident observed, assisted, or produced during a patient care encounter that listed the resident in the health record as a care provider.
(M3) *Resident Productivity* (**RP**) equals assigned-RVUs that residents produced per resident-day. Residents are said to produce workload directly or by assisting staff.
(M4) *Resident Net Productivity* (**RNP**) equals RP minus an RVU opportunity cost that would have been produced in the clinic by attending ophthalmologists had their time to produce care not been diverted to the supervision of residents.
(M5) *Physician Productivity* (**PP**) equals attending physician produced clinic RVUs per attending physician-day.
(M5a) *Average Non-Supervising Physician Productivity* (***avgNS*PP**) equals the VA-wide average PP during a given academic year on clinic-days when no residents were present.

Rates:

(R1) *Workload Assigned Rate* (**WArate**) equals $\frac{[\mathrm{RAW}]}{[avgNS\mathrm{PP}]}$ and compares assigned workload per resident to the average non-supervising productivity rate of VA ophthalmologists.
(R2) *Workload Produced Rate* (**WPrate**) equals $\frac{[\mathrm{RP}]}{[\mathrm{RAW}]}$ and compares the resident produced workload to the total resident assigned-workload. WPrate ranges from 0% (resident only observed) to 100% (resident produced all assigned-workload).
(R3) *Physician Productivity Rate* (**PPrate**) equals $\frac{[\mathrm{PP}]}{[avgNS\mathrm{PP}]}$ and compares the attending ophthalmologist's in-clinic productivity rate to the average non-supervising productivity rate of VA attending ophthalmologists.

Indices:

(I1.) *Supervision Index* (**SI**) equals $\frac{[\mathrm{RNP}]}{[\mathrm{RP}]}$ and assesses the extent residents progressively worked autonomously from supervision when producing workload. SI ranges from 0% (dependence) to 100% (progressive autonomy).
(I2.) *Workload Index* (**WI**) equals $\frac{[\mathrm{RP}]}{[avgNS\mathrm{PP}]}$ and compares the resident productivity rate with the average non-supervising productivity rate of VA attending ophthalmologists.

$$[\mathrm{WI}] = \frac{[\mathrm{RP}]}{[avgNS\mathrm{PP}]} = \frac{[\mathrm{RAW}]}{[avgNS\mathrm{PP}]} \times \frac{[\mathrm{RP}]}{[\mathrm{RAW}]} = [\mathrm{WArate}] \times [\mathrm{WPrate}]. \quad (1)$$

(I3.) *Readiness Index* (**RI**) equals $\frac{[\mathrm{RNP}]}{[avgNS\mathrm{PP}]}$ and compares the resident productivity rate net of supervision with the average non-supervising productivity rate of VA attending ophthalmologists. RI ranges from 0% (not ready) to 100% (practice ready). The Readiness Index comprises both Workload and Supervision Indices.

$$[\mathrm{RI}] = \frac{[\mathrm{RNP}]}{[avgNS\mathrm{PP}]} = \frac{[\mathrm{RAW}] \times \frac{[\mathrm{RP}]}{[\mathrm{RAW}]} \times \frac{[\mathrm{RNP}]}{[\mathrm{RP}]}}{[avgNS\mathrm{PP}]} = \frac{[\mathrm{RAW}]}{[avgNS\mathrm{PP}]} \times \frac{[\mathrm{RP}]}{[\mathrm{RAW}]} \times \frac{[\mathrm{RNP}]}{[\mathrm{RP}]} = [\mathrm{WI}] \times [\mathrm{SI}] \quad (2)$$

Data

Data came from clinic and service-date aggregated, de-identified extracts of VA's electronic health records processed through the Oak Ridge National Laboratory from the U.S. Department of Energy. Additional administrative data on the number of residency positions and facility-level HEDIS measures came from the Veterans Health Administration. The authors are federal employees with the approval of and supervision from their respective agency head to conduct these secondary analyses, subject to the Ethical Framework Principles for Access to and Use of Veteran Data (38 C.F.R. §0.605) overseen by VA's Office of Ethics in Healthcare and are thus exempt from institutional review per 45 C.F.R. §§ 46.104(d)(4)(ii) & 46.101(c).

Inputs and Outputs

Provider and resident counts are computed in clinic-days by specialty, clinic location, and service-date where 2 half-days count as 1 provider-day. Provider specialties are classified by the American Medical Association's National Uniform Claim Committee's Provider Taxonomy code set.[13]

The volume and complexity of clinic workload is quantified in RVUs determined from workload charges from the National Physician Fee Schedule Relative Value Files, with clinic visit and procedure codes classified by Current Procedure Terminology[14] (CPT).

GME Factors

Data about VA ophthalmology residents came from OAA's financial records and VA's electronic patient health records. Resident average academic level was computed over a given academic year as the mean number of PGY01 positions weighted by 0.5 years, PGY02's by 1.5 years, PGY03's by 2.5 years, and PGY04's by 3.5 years. Program size equals the complement of full-time ophthalmology positions VA funds at each of its medical centers, with residents coming from 1 or more programs sponsored by an affiliated education institution and accredited by the Accreditation Council of Graduate Medical Education. Professional staffing was computed as the clinic's resident-to-physician staff ratio, or the number of residents per attending ophthalmologist averaged over half shifts. Patient complexity was computed as RVUs-per-patient totaling all patient care encounters for a given clinic and service-date where a resident had been listed as a care provider in the patient's health record.

GME Covariates

There were no direct measures of outpatient ophthalmology care quality that can be discerned from these patient health records. As a proxy, we computed the medical center's rank of Healthcare Effectiveness Data and Information Set (HEDIS) measures to reflect its culture of emphasizing care quality. Ranks were classified by quartiles, determined for each calendar year based on averaging all 20 HEDIS measures VA medical centers tracked during the study period.[15] The items' mnemonic classifications are listed in Table 1, footnote #d.

VA scores medical center complexity on a 5-point, ordinal scale based on care activities, services offered, training programs, and research activities.[16] Medical centers were classified as a "main hospital" or "network facility" within the main hospital's network of facilities. The standard deviation of RVU-per-patient treated was computed quarterly across service-dates by clinic.

Analyses

Details regarding how metrics were computed are described in the Appendix. Overall, productivity rates were estimated using Generalized Linear Mixed Models (GLMM)[17,18] where RVU workload was regressed on resident and provider counts, with random coefficients computed by clinic at 2-month intervals for resident and attending physician counts plus their 2-way interaction. Model selection was based on Best Approximating Model methods (BAM)[19,20] with a Generalized Bayesian Information Criterion (GBIC).[21] The BAM method searched over all possible combinations of main effects variables in a covariate pool comprising resident and provider counts plus statistically significant ($p = .05$) 2-way interactions and second-order polynomials using the robust sandwich estimator.[20-22] The best models were selected by applying an Occam's Window.[23] Final estimates were computed by averaging estimates from BAMs included in the Occam's Window weighted by their respective GBIC fit. Interpreting estimates of the model's parameter coefficients to estimate resident productivity net of supervision was previously explained mathematically and tested empirically.[7]

Metric, rates, and indices were described by means, standard deviations (SD), interdecile ranges (ID) and intraclass correlation coefficients (ICC). Independent associations between indices and GME factors were computed using GLMM with GME factors as independent variables. Estimates of association size are exploratory measures to reveal pattern and thus are not tests of predetermined hypotheses.

RESULTS

Sample

A total 246 VA facilities reported 153,415 clinic-days with VA's ophthalmology clinic code during academic years from July 1, 2015, through June 30, 2019. Of these, 140,585 days from 228 facilities met criteria for a clinic-day and 95,817 days from 110 medical centers met criteria for a teaching clinic-day. The sample was reduced by 1 clinic that did not have a facility complexity score, leaving 94,847 clinic-days and 109 facilities. Over the 4 study years, VA budgeted 1,327.24 annual ophthalmology resident positions. Residents filling those positions rotated through 86,530 clinic-days of which 84,593 (97.8%) included only PGY02-04 ophthalmology residents assigned to 2,347,452 (38.7%) of 6,062,944 patient encounters and assigned to a total 8,837,993 (57.6%) of 15,352,202 RVUs.

WRAP Measures

Displayed in Tables 1-3, teaching clinics averaged 82 RVUs/resident (ID:25-162) of which residents were assigned 39 RVUs/resident (ID:14-64) and produced 30 RVUs/resident (ID:10-51) representing 79% of their assigned RVUs, and produced net of supervision 21 RVUs/resident (ID:4-39) representing 66% (ID:28-100) of resident produced workload and 51% (ID:9-97)

TABLE 1. Sample Size, Hospital Characteristics, Clinic Output, and Number of Providers per Clinic, by Academic Year

Category	Symbol	All Years[a]	2016[a]	2017[a]	2018[a]	2019[a]
Sample size						
Number of Hospitals		109	97	96	97	98
with residents		109	96	95	96	97
Number of Clinic-days		94,847 (100.0%)	23,765 (100.0%)	23,291 (100.0%)	23,640 (100.0%)	24,151 (100.0%)
with residents		84,593 (89.2%)	21,251 (89.4%)	20,936 (89.9%)	21,064 (89.1%)	21,342 (88.4%)
Workload Output (clinic-day)						
Patients Treated	(Pt)	66.2 (42.3)	66.6 (43.4)	66.8 (42.3)	65.3 (41.2)	65.9 (42.5)
resident assigned	(Pt^R)	24.9 (19.0)	25.6 (19.4)	25.5 (18.9)	24.6 (18.4)	24.1 (19.1)
Procedures (RVU)	(W)	169.9 (125.0)	161.5 (119.8)	165.3 (121.3)	172.1 (123.4)	180.6 (134.0)
resident Assigned	(W^R)	93.6 (81.8)	92.6 (80.4)	91.9 (76.4)	94.5 (79.3)	95.5 (85.4)
Provider Mix (clinic-day)[b]						
Physician	(PHY)					
Number days present		94,847 (100.0%)	23,765 (100.0%)	23,291 (100.0%)	23,640 (100.0%)	24,151 (100.0%)
Size when present		2.63 (1.85)	2.59 (1.75)	2.64 (1.87)	2.66 (1.94)	2.64 (1.84)
Physician Resident	(PHYR)					
Number days present		84,593 (89.2%)	21,251 (89.4%)	20,936 (89.9%)	21,064 (89.1%)	21,342 (88.4%)
Size when present		2.73 (1.59)	2.72 (1.55)	2.73 (1.55)	2.74 (1.62)	2.72 (2.75)
Other Clinician	(OTH)					
Number days present		73,803 (77.8%)	18,614 (78.3%)	18,329 (78.7%)	18,426 (77.9%)	18,434 (76.3%)
Size when present		3.04 (2.45)	2.89 (2.40)	3.00 (2.41)	3.03 (2.41)	3.22 (2.55)
Nurse RN	(RN)					
Number days present		15,324 (16.2%)	3,576 (15.1%)	3,773 (16.2%)	3,946 (16.7%)	4,029 (16.7%)
Size when present		0.76 (0.46)	0.70 (0.34)	0.84 (0.59)	0.75 (0.44)	0.75 (0.41)
Physician Assistant	(PA)					
Number days present		8,484 (8.9%)	2,210 (9.3%)	2,302 (10.0%)	2,030 (8.6%)	1,942 (8.0%)
Size when present		0.71 (0.25)	0.69 (0.24)	0.72 (0.25)	0.73 (0.26)	0.70 (0.25)
Nurse Practitioner	(NP)					
Number days present		7,316 (7.7%)	1,631 (6.9%)	1,767 (7.6%)	1,756 (7.4%)	2,162 (9.0%)
Size when present		0.84 (0.43)	0.82 (0.38)	0.89 (0.48)	0.85 (0.45)	0.79 (0.50)
Pharmacist	(PHARM)					
Number days present		368 (0.4%)	81 (0.3%)	77 (0.3%)	115 (0.5%)	95 (0.4%)
Size when present		0.51 (0.05)	0.52 (0.10)	0.50 (0.00)	0.50 (0.05)	0.50 (0.00)
Psychologist	(PSY)					
Number days present		217 (0.2%)	65 (0.3%)	43 (0.2%)	55 (0.2%)	54 (0.2%)
Size when present			0.50 (0.00)	0.50 (0.00)	0.51 (0.01)	0.50 (0.00)
Other Clinician – Trainee	(OTHT)					
Number days present		463 (0.5%)	51 (0.2%)	14 (0.1%)	141 (0.6%)	257 (1.1%)
Size when present		0.78 (0.34)	0.60 (0.20)	0.54 (0.13)	0.77 (0.30)	0.84 (0.37)
GME Factors						
Average PGY level	(PGYavg)	2.68 (0.39)	2.68 (0.39)	2.69 (0.38)	2.65 (0.40)	2.69 (0.39)
Percent at PGY02	(PGY02)	26.17%	26.81%	24.78%	27.30%	25.79%
Percent at PGY03	(PGY03)	29.91%	28.62%	31.75%	29.96%	29.33%
Percent at PGY04	(PGY04)	43.92%	44.57%	43.47%	42.74%	44.88%
Residency Program Size[c]	(PGYn)	3.77 (1.58)	3.72 (1.58)	3.81 (1.55)	3.71 (1.57)	3.83 (1.61)
RVU/patient	(W/Pt)	2.84 (3.28)	2.64 (2.37)	2.60 (1.93)	2.98 (3.64)	3.12 (4.45)
Resident Assigned Care	(W^R/Pt^R)	3.93 (4.64)	3.66 (2.74)	3.64 (3.14)	4.17 (5.87)	4.24 (5.83)
Resid–to–Phy Ratio	(RSR)	1.17 (1.10)	1.15 (1.03)	1.17 (1.03)	1.18 (1.16)	1.18 (1.19)
GME Covariates						
HEDIS Quality[d]	(Q)	86.6 (2.0)	86.9 (2.1)	86.5 (2.0)	86.5 (1.9)	86.5 (2.0)
Std. Dev. RVU/patient[e]	σ(W/Pt)	1.61 (3.50)	1.37 (2.33)	1.25 (2.15)	1.93 (4.62)	1.90 (4.12)
Resident Assigned Care	σ (W^R/Pt^R)	2.32 (4.22)	2.02 (2.42)	1.90 (2.86)	2.72 (5.72)	2.64 (4.86)
Hospital Status						
Main Hospital		81,459 (85.9%)	21,121 (88.9%)	20,503 (88.0%)	19,747 (83.5%)	20,088 (83.2%)
Network		13,388 (14.1%)	2,644 (11.1%)	2,788 (12.2%)	3,893 (16.5%)	4,063 (16.8%)
Hospital Complexity[f]	(HComp)					
1a–High		46,898 (49.4%)	11,664 (49.1%)	11,463 (49.2%)	11,868 (50.2%)	11,903 (49.3%)
1b–High		21,074 (22.2%)	5,023 (21.1%)	5,057 (21.7%)	5,239 (22.2%)	5,755 (23.8%)
1c–High		19,210 (20.3%)	5,040 (21.1%)	4,732 (20.3%)	4,699 (19.9%)	4,739 (19.6%)

(continued)

TABLE 1 (*continued*)

Category	Symbol	All Years[a]	2016[a]	2017[a]	2018[a]	2019[a]
2—Medium		5,222 (5.5%)	1,427 (6.0%)	1,435 (6.2%)	1,195 (5.1%)	1,165 (4.8%)
3—Low		2,443 (2.6%)	611 (2.6%)	604 (2.6%)	639 (2.7%)	589 (2.4%)

[a]Counts are reported as numbers with percentages in parentheses (%). Continuous measures are reported as means with standard deviation in parentheses. Values are computed by clinic-day (pertaining to a clinic on a given service-date), clinic-year (pertaining to a clinic over an academic year), facility-year (pertaining to a facility over an academic year), and facility (pertaining to a facility over the 4 study academic years).
[b]Number of clinic-days when a provider by specialty is present and assigned to patient care, and the mean and standard deviation in the number of providers present and assigned to patient care during a clinic-day. Excluded were provider specialties where there were fewer than 5 clinic-days per year.
[c]Information on residency positions by medical center and academic year was obtained from VA's Office of Academic Affiliations financial records specifying the number of full-time residency positions in ophthalmology, by PGY level and academic year. During 2016-2019 academic years, a total 1,327.24 yearly, full-time ophthalmology residency positions were funded to VA medical centers, with 27.65 (2.08%) PGY01 positions and 357.56 (26.94%) PGY02's, 384.01 (28.93%) PGY03's, and 558.02 (42.04%) PGY04's. There were no allocated PGY05 or higher positions. PGY01 residents were excluded from analyses because of their small numbers and lack of meaningful contributions to workload. The number of PGY positions (PGYn) was obtained by summing the number of PGY02, PGY03, and PGY04 positions. The average PGY level (PGYavg) by medical center and academic year was obtained by multiplying 1.5 years to the number of PGY2 positions, multiplying 2.5 to PGY3's, and 3.5 to PGY4's, summing the products, and dividing the total by the total positions funded (PGYn). (PGY02), (PGY03) and (PGY04) are the respective proportions of PGY02, PGY03, and PGY04 positions to the total funded (PGYn) by medical center by academic year. Finally, the distribution of residents by PGY level for network facilities hosting residents in 13,388 (14.1%) of 94,847 clinic-days were inferred from PGY levels funded at the respective main hospital.
[d]Facility-level HEDIS Quality metric is computed by calendar year as the average of 20 Healthcare Effectiveness Data and Information Set (HEDIS) measures that VA reported during the time spanning the 4 academic years.Comprehensive Diabetes Care: Percent of diabetes outpatients with (c9h) HbA1c testing from medical record abstracts; (dmg23h) poor HbA1c control (reversed scored) from medical record abstraction, electronic measurement, and external peer review program; (dmg27h) blood pressure control (<140/90) from medical record abstraction, electronic measurement, and external peer review program; (dmg31h) eye exams from medical record abstraction; (dmg34h) medical attention for nephropathy from medical record abstraction, electronic measurement, and external peer review program.Cardiovascular Health: (ihd20h) percent of discharges for patients with Acute Myocardial Infarction receiving persistent Beta-Blocker treatment from medical record abstraction.Mental Health: Percent of patients: (mdd40) with major depression receiving an annual screen for depression from medical record abstraction; (ptsd51) screened for PTSD from medical record abstraction; (sa17) with alcohol screen who had a timely intervention from medical record abstraction; (sa7) with alcohol misuse screen from medical record abstraction; and (sre1) with positive PTSD/MDD screen with timely Suicide Risk Evaluation from medical record abstraction.General Primary Care and Preventive Services: (p25h) percent of patients age 65 and older receiving flu vaccinations from medical record abstraction; (p26h) percent of patients ages 18-64 receiving flu vaccinations from medical record abstraction, electronic measurement, and external Peer Review Program; (p61h) patients 50-75 with colorectal cancer screen from medical record abstraction; and patients who smoke and (smg8) were advised to quit from VA medical record abstraction; (smg9) had discussions on cessation strategies from VA medical record abstraction; and (smg10) had discussion on cessation medications from VA medical record abstraction.Women's Health: Percent of women (p41h) age 21-64 receiving cervical cancer screening from medical records abstraction; (p42) age 21-29 receiving cervical cancer screening from medical record abstraction; (p43h) age 30-64 receiving cervical cancer screening from medical record abstraction.HEDIS measures that are not available for a network facility is inferred from its corresponding main hospital.
[e]RVU/patient treated is calculated for each clinic-day. The standard deviation of RVU/patient treated is computed for each clinic over an academic quarter.
[f]5-item ordinal hospital complexity scale is conducted by VA Office of Productivity, Efficiency and Staffing every 3-5 years and reflects mix of clinical services, research produced, and education and training activities.

of the average attending productivity rate on nonteaching clinic-days. WRAP measures clustered by clinic and academic year (ICCs from 62% to 88%) and by clinic and its 4-year clinic average (ICCs from 46% to 81%).

There was minimal correlation ($r = 0.11 \pm 0.01$) between SI and WI as components of the Readiness Index, and between WPrate and WArate ($r = -0.14 \pm 0.01$) as components of the Workload Index.

Independent Associations

Table 4 displays estimates and 95% confidence intervals in brackets of the independent associations between GME factors and selected WRAP indices and rates.

Clinics where residents had advanced by 1 PGY level saw an average 4.1% (95% CI, [3.4, 4.7)) increase in readiness as residents had more progressive autonomy from supervision (5.0%, 95% CI, [4.4, 5.4]). There was, however, no change in the resident workload rate as higher assigned rates (3.2%, 95% CI, [2.1, 4.2]) were offset by lower resident produced rates (−2.8%, 95% CI, [−3.1, −2.4]).

For clinics hosting 1 additional resident per attending ophthalmologist, readiness increased by 5.6% (95% CI, [5.3, 5.9]) as attending ophthalmologists entrusted their residents with more progressive autonomy from supervision (3.9%, 95% CI, [3.8, 4.1]) and residents produced more workload as clinics shifted production away from attending ophthalmologists (−13.7%, 95% CI, [−14.2, −13.3]) to residents (4.2%, 95% CI, [3.8, 4.7]).

For clinics whose patient mix added 1 RVU per-patient-treated, readiness increased by only 0.7% (95% CI, [0.5, 0.9]). The small increase was due to the increase in workload (1.4%, 95% CI, [1.1, 1.8]) offset in

TABLE 2. Means, Percentiles, and Correlations of VA's Workload-based Resident Academic Performance Metrics for Ophthalmology Residents in VA Outpatient Ophthalmology Clinics Computed at the Clinic and Service-Date Level

	Resident Clinic Workload (RCW) $(\frac{\text{RVU}}{\text{resident day}})$	**Resident Assigned Workload [RAW]** $(\frac{\text{RVU}}{\text{resident day}})$	**Resident Productivity [RP]** $(\frac{\text{RVU}}{\text{resident day}})$	**Resident *Net* Productivity [RNP]** $(\frac{\text{RVU}}{\text{resident day}})$	**Att. Phy. Productivity [PP]** $(\frac{\text{RVU}}{\text{physician day}})$
Mean	81.94	38.49	29.84	20.67	22.28
Standard Deviation	73.61	23.42	19.34	15.01	20.13
Percentiles					
10th %	25.28	14.22	10.37	3.79	1.74
50th %	60.21	35.25	27.56	18.78	20.03
90th %	162.25	63.79	50.60	39.13	45.49
ICC[a] *with*					
Clinic 4-yr. avg	0.74	0.56	0.50	0.58	0.81
Clinic annual avg	0.79	0.67	0.62	0.70	0.88
BAM Estimates[b]	*	*	1.00	1.00	1.00
Correlation *with*[c]					
RCW		0.41	0.22	0.26	0.59
RAW			0.82	0.68	0.47
RP				0.73	0.35
RNP					0.31

[a]Intraclass Correlations were based on 2-way mixed effects models, absolute agreement, and average measures, with 95% confidence intervals within ±0.01 of the reported value. All Intraclass Correlations were significant at the .001 level.
[b]BAM Intraclass Correlations compare the respective model average with individual BAM estimates. All Intraclass Correlations were significant at the .001 level.
[c]Pearson Correlations were statistically significant at the .001 level, with 95% confidence interval within ±0.01 of the reported value.

TABLE 3. Means, Percentiles, and Correlations of Selected WRAP Indices and Rates for Ophthalmology Residents in VA Teaching Outpatient Ophthalmology Clinics Computed at the Clinic and Service-Date Level

	Readiness Index [RI] $(\frac{\text{RNP}}{avgNS\text{PP}})$	**Supervision Index [SI]** $(\frac{\text{RNP}}{\text{RP}})$	**Workload Index [WI]** $(\frac{\text{RP}}{avgNS\text{PP}})$	**Workload Produced Rate [WPrate]** $(\frac{\text{RP}}{\text{RAW}})$	**Workload Assigned Rate [WArate]** $(\frac{\text{RAW}}{avgNS\text{PP}})$	**Att. Phy. Productivity Rate [PP]** $(\frac{\text{PP}}{avgNS\text{PP}})$
Mean	51.01%	65.71%	73.65%	78.93%	95.00%	55.00%
Standard Deviation	37.06%	26.81%	47.74%	15.57%	57.81%	49.67%
Percentiles						
10th %	9.36%	26.11%	25.60%	57.73%	35.10%	4.31%
50th %	46.35%	69.05%	68.01%	81.26%	87.00%	49.44%
90th %	96.59%	100.00%	124.89%	98.06%	157.44%	112.29%
ICC[a] *with*						
Clinic 4-yr. avg	0.58	0.46	0.50	0.72	0.56	0.81
Clinic annual avg	0.70	0.63	0.62	0.84	0.67	0.88
BAM estimates[b]	1.00	1.00	1.00	1.00	*	1.00
Correlation *with*[c]						
Supervision Index	0.58					
Workload Index	0.73	0.11				
Workload Produced Rate	0.25	0.13	0.21			
Workload Assigned Rate	0.68	0.05	0.82	-0.14		
Physician Productivity Rate	0.31	0.13	0.35	-0.21	0.47	

[a]Intraclass Correlations were based on 2-way mixed effects models, absolute agreement, and average measures, with 95% confidence intervals within ±0.01 of the reported value. All Intraclass Correlations were significant at the .001 level.
[b]BAM Intraclass Correlations compare the respective model average with individual BAM estimates. All Intraclass Correlations were significant at the .001 level.
[c]Pearson Correlations were statistically significant at the .001 level, with 95% confidence interval within ±0.01of the reported value.

TABLE 4. Independent Associations[a] Between GME Factors and Selected WRAP Indices and Rates for Ophthalmology Residents in VA Teaching Outpatient Ophthalmology Clinics

	Readiness Index (%)	Supervision Index (%)	Workload Index (%)	Workload Produced Rate (%)	Workload Assigned Rate (%)	Att. Phy. Prod. Rate (%)
Adjusted Means[e]	47.109 (46.269, 47.948)	60.500 (59.802, 61.198)	72.989 (72.015, 73.963)	81.395 (81.033, 81.757)	92.139 (90.814, 93.463)	58.564 (57.325, 59.803)
GME Factors						
Academic Level (yrs.)	4.059 (3.417, 4.702)	4.913 (4.378, 5.448)	0.479 (-0.291.1.249)[b]	-2.789 (-3.072, -2.506)	3.178 (2.145, 4.210)	-1.217 (-2.150, -0.284)[c]
Resid—to—Phy Ratio	5.556 (5.255, 5.858)	3.935 (3.770, 4.100)	4.213 (3.771, 4.654)	4.576 (4.458, 4.693)	-0.683 (-1.178, -0.189)[d]	-13.744 (-14.212, -13.276)
RVU/patient-treated	0.698 (0.509, 0.888)	-0.160 (-0.200, -0.119)	1.420 [1.067. 1.773)	-0.133 (-0.158, -0.108)	2.147 (1.655, 2.640)	1.599 (1.292. 1.906)
Resident Program Size	0.393 (0.198, 0.588)	-0.743 (-0.885, -0.602)	0.935 (0.634, 1.235)	-0.086 (-0.165, -0.008)[c]	1.060 (0.786, 1.333)	-3.619 (-3.857, -3.380)
Adjusting Covariates[f]						
HEDIS quartile rank	-1.792 (-2.024, -1.561)	-1.395 (-1.568, -1.223)	0.211 (-0.086, 0.508)[b]	-0.186 (-0.281, -0.091)	0.635 (0.307, 0.962)	1.472 (1.147, 1.796)
Std. Dev. RVU/patient	-0.141 (-0.194, -0.088)	-0.087 (-0.119, -0.056)	-0.184 (-0.253, -0.116)	-0.095 (-0.113, -0.077)	-0.149 (-0.230, -0.068)	0.360 (0.293. 0.428)
Annual Progress (qtrs.)	0.349 (0.148, 0.549)	-0.411 (-0.563 -0.259)	1.039 (0.756, 1.322)	-0.111 (-0.158, -0.108)[c]	1.206 (0.917, 1.494)	0.441 (0.192, 0.690)
Attending Productivity	0.608 (0.586, 0.630)	0.212 (0.200, 0.224)	0.821 (0.788, 0.853)	-0.096 (-0.102, -0.091)	1.275 (1.229, 1.321)	*
Resident Productivity	*	*	*	*	*	0.979 (0.942, 1.016)
Network Status	-10.234 (-11.040, -9.428)	-4.053 (-4.649, -3.458)	-7.373 (-8.266, -6.481)	-0.759 (-1.060, -0.458)	-9.301 (-10.436, -8.166)	3.656 (2.713, 4.598)
Hospital Complexity:						
High—1a (referent)						
High—1b	-1.970 (-2.520, -1.420)	-0.700 (-1.165, -0.234)	-0.831 (-1.448, -0.214)[d]	1.466 (1.221, 1.711)	-1.825 (-2.598, -1.052)	-11.648 (-12.328, -10.969)
High—1c	3.956 (3.271, 4.641)	-0.800 (-1.306, -0.293)	7.089 (6.252, 7.926)	0.855 (0.576, 1.134)	7.496 (6.429, 8.564)	3.363 (2.416, 4.310)
Medium—2	-9.135 (-10.136, -8.134)	1.625 (0.797, 2.454)	-10.701 (-12.009, -9.392)	4.058 (3.652, 4.464)	-18.261 (-19.956, -16.566)	-4.598 (-6.025, -3.170)
Low—3	11.772 (10.084, 13.461)	3.320 (2.391, 4.249)	12.943 (10.565, 15.321)	7.458 (6.910, 8.007)	4.220 (1.286, 7.155)[d]	14.020 (12.216, 15.824)

[a]Independent associations are descriptive and computed using a Generalized Linear Model. The Wald $\chi^2(1)$ distributed 95 percent confidence intervals of parameter estimates are presented in brackets. All values are statistically significant at the <.001 level unless otherwise specified. Values are reported as RVUs/resident-day, RVUs/attending physician-days, or percent.
[b]Value is not significant at the .05 level.
[c]Value is significant at the .05 level.
[d]Value is significant at the .01 level.
[e]Baseline is the expected value when all independent continuous factors are set at their mean value, with referent main hospitals and the highest complexity level facilities (high -1a).
[f]Covariate coefficient estimates are provided for completeness, but as adjusters should not be interpreted.

part by a decrease in resident progressive autonomy from supervision (−0.16%, 95% CI, [−0.20, −0.11]).

For each position added to the complement of ophthalmology residency positions rotating through a VA clinic, readiness increased by only 0.4% (95% CI, [0.2, 0.6]). The small increase was due to the increase in workload, a result of clinics shifting production away from attending ophthalmologists (−3.6%, 95% CI, [−3.9, −3.4]) to residents (0.9%, 95% CI, [0.6, 1.2]), and offset in part by a decrease in resident progressive autonomy from supervision (−0.7%, 95% CI, [−0.9, −0.6]).

DISCUSSION

This study introduced VA's new WRAP measures and demonstrated how resident academic level, program size, professional staffing, and patient complexity of GME teaching ophthalmology clinics may impact ophthalmology residents' preparedness to meet the clinical workload demands of independent practice. Overall, the Readiness Index averaged 51% (ID[9%-97%]) across VA ophthalmology clinics and service-dates between July 1, 2015, and June 30, 2019. This means the average ophthalmology resident productivity net of supervision to produce RVUs in outpatient clinic encounters was estimated at 51% of the average nonsupervising productivity rate among VA ophthalmologists during the same period. This compares to an earlier estimate of 54% ID (10%-98%) computed for all VA surgery residents engaged in outpatient care between 2005 and 2018.[7]

Consistent with the progressive independence hypothesis,[8-10] these data revealed more experienced residents (higher PGY level) had been entrusted with more progressive autonomy from supervision and were thus better prepared to meet the workload demands of independent practice than their less experienced counterparts. Contrary to expectation, however, resident productivity rates were no higher even though senior residents were assigned to more workload. One explanation is senior residents may have been intentionally exposed to diverse groups of patient encounters to expand their clinical learning experience. In fact, among teaching clinic-days, 40% involved residents whose assigned workload per resident exceeded the nonsupervising productivity of the average nonsupervising ophthalmologist in the same period.

These data also suggested ophthalmology program directors may improve their residents' workload readiness by increasing the number of residents per supervising physician. For example, if an average VA clinic with 1.2 residents per physician were to have been operating at 1 standard deviation higher to 2.3 residents per physician, the readiness of the average resident may increase the equivalent of having had an additional 18.1 months of clinical experience. The greater readiness is expected to come from residents being entrusted with more progressive autonomy while producing more workload as clinics shifted production from attending ophthalmologists to residents. While such a shift in workload allowed attending ophthalmologists more time for direct supervision, the increase in available supervisory time was offset by the increase in the number of residents each attending ophthalmologist was responsible to supervise. This pattern of attending staff shifting workload to their residents is consistent with the current literature reporting attending physicians with more GME supervising responsibilities tended to have lower patient care productivity rates.[24-26]

Program directors may also improve resident workload readiness by assigning residents to more complex patients. For example, if the average VA clinic where residents had been assigned 3.9 RVUs per-patient-treated were to have been operating by 1 standard deviation higher at 8.6 RVUs/patient, the readiness of the average resident may increase the equivalent of having had an additional 9.6 months of clinical experience. The increase in readiness was a likely result of increasing resident and attending ophthalmologist productivity rates in response to the greater demands being placed on clinic resources to manage more complex patient cases. On the other hand, the increase in readiness was moderated by residents being entrusted with less progressive autonomy from supervision because their patients presented with more complex care needs.

Finally, program directors may also improve resident readiness by expanding program size. For example, if the average clinic supporting a total 3.8 residents were to have been supporting 1 standard deviation higher at 5.3 residents, the readiness of the average resident may increase the equivalent of having had an additional 1.8 months of clinical experience. The increase was the result of attending ophthalmologists shifting workload from themselves to the growing number of residents while on the other hand entrusting each resident with less progressive autonomy from supervision. This is inconsistent with studies of general surgery in the operating room where a larger complement of residents was associated with more supervision autonomy,[27] though such findings may reflect residents having been "over-supervised".[28]

The application of WRAP measures to GME evaluation studies offers several advantages. Indices are scored at the clinic and service-date level from objective patient health records. Index values fall along continuous scales based on RVUs. As a measure of workload, RVUs are well-known among health professionals. RI is computed from production models selected following exhaustive

searches of all plausible input variables, with final estimates computed by model averaging to account for both sampling errors and model uncertainty.[19,20] RI incorporates both WI and SI to reflect what residents do and the supervision they received when doing it. This is consistent with Terhune[29] suggestion that academic medical centers should offer policy makers and the public 'transparency' about their residents and the supervised care they engage as part of their training.

There is an empirical basis for assessing resident readiness in terms of resident workload and its supervision. The number of patients seen and billable charges have both been linked to care quality,[30,31] resident education outcomes,[32,33] and patient health outcomes.[34] The intensity of resident supervision has been linked to patient care outcomes,[4] to resident workload production[7] and progressive independence, and to the volume and complexity of patient care.[8-10] Overall, the WRAP continuous-scaled, health record-computed, and objective-based Supervision Index offers a viable alternative to traditional measures of resident supervision including those based on perceived adequacy,[35-38] entrusted levels,[39-42] or classifications of attending physicians' time.[10] Such mostly ordinal scales have had mixed reviews with some reporting correlations with resident Post Graduate Year (PGY),[43] patient[44,45] and education[35,36,44,45] outcomes, and care quality,[36] while others reporting no correlations with outcomes in outpatient ophthalmology clinics,[37,38] inpatient general-medical service,[46] surgery operations,[3,47-49] and pediatric emergency departments.[50] These inconsistencies, alternatively, may be the result of small sample sizes and from quantifying supervision levels from resident and attending physician survey data whose administration may disrupt clinic workflow,[43] responses that are subject to perception biases,[51] and ordinal scores that fail to capture resident performance.

There are limitations to these findings. First, this study only focused on ophthalmology resident productivity. There are other competencies to consider when determining whether a resident is prepared to enter independent practice. Second, our focus on VA ophthalmology may not generalize to other surgery specialties or to non-VA facilities.

Third, facility-level HEDIS measures were used as a proxy covariate to control for variation care quality. This is necessary if higher productivity rates were achieved at the expense of care quality. For these data, we found residents in facilities with a higher quartile HEDIS rank tended to be entrusted with less progressive autonomy from supervision (−1.4%, 95% CI, [−1.6, −1.2]) and readiness to enter independent practice (−1.8, 95% CI, [−2.0, −1.6]). However, the results of this study must be interpreted with caution as these facility-level HEDIS measures may not have captured all variation in care quality. On the other hand, VA has attempted to standardize its care processes as revealed by only a 2.3% coefficient of variation in HEDIS scores across VA facilities by calendar year. VA strictly enforces adherence to its policies governing resident supervision.

Fourth, RVUs were not categorized by subspecialty within ophthalmology. Certain subspecialties, such as neuro-ophthalmology, tend to have more complicated patients and longer examination times than others with the same number of RVUs. Fifth, coding errors to list residents as care providers in the health record may bias estimates of WRAP measures. Finally, resident PGY levels were inferred from VA's allocation of paid specialty positions by PGY level to VA facilities by academic year.

CONCLUSION

Academic medicine policy makers should consider the roles that GME clinic factors play to progress residents towards independent practice. When assessing academic medical centers, evaluators should add WRAP Readiness, Supervision, and Workload Indices to traditional measures of patient health outcomes, care quality, resident satisfaction and resident education milestone achievement.

ACKNOWLEDGMENTS

This study was made possible with the support from the U.S. Veterans Health Administration, Office of Academic Affiliations and from the Office of Research and Development, Health Services Research and Development Service (SDR#13-426, IIR#14-071, IIR#15-084).

This study also used resources of the Knowledge Discovery Infrastructure at the Oak Ridge National Laboratory supported by the Department of Energy under Contract No. DE-AC05-00OR22725 and funded by the Department of Veterans Affairs Inter-Agency Agreement and by VA-DoD Joint Incentive Fund under IAA No. 2450-Z127-17. Programming, data access, and administrative support was provided by Ian Goethert, Hope Cook, and Laura Davies through the Oak Ridge National Laboratory in Oak Ridge, TN, and Merry Ward, Health Informatics Research and Innovations Manager, Office of Health Informatics, Veterans Health Administration, Department of Veterans Affairs.

We also acknowledge contributions from Cristian Peralta, Chief Financial Officer, Washington, D.C. and David S. Bernett, Director, Data Management and Support Center, St. Louis, MO in the Office of Academic Affiliations, US Department of Veterans Affairs, for their support for this study.

REFERENCES

1. Franzone JM, Kennedy BC, Merritt HM, Casey JT, Austin MC, Daskivich TJ. Progressive independence in clinical training: perspectives of a national multispecialty panel of residents and fellows. *J Grad Med Ed*. 2015;7(4):700–704. https://doi.org/10.4300/JGME-07-04-51.
2. Bell RH. Why johnny cannot operate. *Surgery*. 2009;146:533–542.
3. Kunac A, Oliver JB, McFarlane JL, Anjaria DJ. General surgical resident operative autonomy vs patient outcomes: are we compromising training without net benefit to hospitals or patients? *J Surg Ed*. 2021;78(6):e174–e182. https://doi.org/10.1016/j.jsurg 2021.09.017.
4. Kashner TM, Greenberg PB, Birnbaum AD, et al. Patient surgical outcomes when surgery residents are the primary surgeon by intensity of surgical attending supervision in veterans affairs medical centers. *Annals of Surgery Open*. 2023;4(4):e3 51. https://doi.org/10.1097/AS9.0000000000000351.
5. I.H.S. Market Ltd. The Complexities of Physician Supply and Demand: Projections from 2019 to 2034. Washington, DC: AAMC; 2021. Accessed March 7, 2024 at www.aamc.org/media/54681/download.
6. Berkowitz SJ, Finn AP, Parikh R, Kuriyan AE, Patel S. Ophthalmology workforce projections in the United States: 2020-2035. *Ophthalmology*. 2024;131(2):133–139. https://doi.org/10.1016/j.ophtha.2023.09.018. S0161-6420(23)00677-2.
7. Kashner TM, Greenberg PB, Henley SS, Bowman MA, Sanders KM. Assessing physician resident contributions to outpatient clinical workload. *Med Care*. 2022;60(9):709–717.
8. Kashner TM, Byrne JM, Henley SS, et al. Measuring progressive independence with the resident supervision index: theoretical approach. *JGME*. 2010;2(1):8–16.
9. Kashner TM, Byrne JM, Chang BK, et al. Measuring progressive independence with the resident supervision index: empirical approach. *JGME*. 2010;2(1):17–30.
10. Byrne JM, Kashner TM, Gilman SC, et al. Measuring the intensity of resident supervision in the department of veterans affairs: the resident supervision index. *Acad Med*. 2010;85(7):1171–1181.
11. Heisler EJ, Mendez BHP, Mitchell A, Panangala SV, Villagrana MA. Federal support for graduate medical education: an overview (R44376). congressional research service, December 27, 2018. https://crsreports.congress.gov (accessed May 23, 2021).
12. Bernett DS, Kashner TM. Annual Report of the Health Services Training Survey Findings for Academic Year 2019-2020 (VHA-OAA Report#0001). Washington, DC: Office of Academic Affiliations, Department of Veterans Affairs; 2021 https://www.va.gov/oaa.
13. National Uniform Claim Committee (NUCC). Health Care Provider Taxonomy, v20.0. American Medical Association; 2020. https://nucc.org/images/stories/PDF/Taxonomy_20_0.pdf. Accessed May 7, 2020.
14. American Medical Association. CPT New Codes. https://www.ama-assn.org/topi9cs/cpt-new-codes. Accessed June 14, 2020.
15. National Committee for Quality Assurance. HEDIS Measures and Technical Resources. accessed https://www.ncqa.org/hedis/measures accessed October 20, 2022.
16. Veterans Health Administration (VHA), Office of productivity, efficiency, and staffing (OPES). veterans health administration (VHA) facility complexity model. http://opes.vassc.med.va.gov/Pages/Facility-Complexity-Model.aspx (accessed March 30, 2021).
17. Stroup WW. Generalized linear mixed models: modern concepts, methods and applications. Boca Raton, FL: CRC Press; 2012.
18. McCulloch CE, Searle SR. Generalized, linear, and mixed models. Hoboken, NJ: John Wiley & Sons; 2004.
19. Kashner TM, Henley SS, Golden RM, Zhou X-H. Making causal inferences about treatment effect sizes from observational datasets. *Biostat Epidemiol*. 2020;4(1):48–83. https://doi.org/10.1080/24709360.2019.1681211.
20. Henley SS, Golden RM, Kashner TM. Statistical modeling methods: challenges and strategies. *Biostat Epidemiol*. 2020;41(1):105–139.
21. Lv J, Liu JS. Model selection principles in Misspecified models. *J Royal Stat Soc: Ser B (Stat Methodol)*. 2014;76(1):141–167.
22. White H. Maximum likelihood estimation of Misspecified models. *Econometrica*. 1982;50(1):1–25.
23. Onorante L, Raftery AE. Dynamic model averaging in large model spaces using dynamic Occam's window. *Eur Eco Rev*. 2016;81:2–14.

24. Garg ML, Boero JF, Christansen RG, Booher CG. Primary care teaching physicians' losses of productivity and revenue at three ambulatory-care centers. *Acad Med*. 1991;66:348–353.

25. Johnson T, Shah M, Rechner J, King G. Evaluating the effect of resident involvement on physician productivity in an academic general internal medicine practice. *Acad Med*. 2008;83(7):670–674.

26. Jones TF, Culpepper L, Shea C. Analysis of the cost of training residents in a community health center. *Acad Med*. 1995;70(6):523–531.

27. Yu Y, Kunac A, Oliver JB, Sehat AJ, DJ Anjaria. General surgery resident complement and operative autonomy-size matters. *J Surg Ed*. 2022;79(6):e76–e84. https://doi.org/10.1016/j.jsurg.2022.09.008.

28. Williams A, Meister R, Crook TW, Swan RR. Resident autonomy: are attendings around too much? *Acad Pediat*. 2019;19(6):e36–e37. https://doi.org/10.1016/j.acap.2019.05.093.

29. Terhune K. Supervision and transparency in resident training (invited commentary). *JAMA Surg*. 2023;158(4):367. https://doi.org/10.1001/jamasurg.2022.7513.

30. Vidyarthi AR, Auerbach AD, Wachter RM, Katz PP. The impact of duty hours on resident self-reports of errors. *J Gen Intern Med*. 2007;22(2):205–209. https://doi.org/10.1007/s11606-006-0065-4.

31. Colt MH, Katz JT, McMahon GT. The effect of workload reduction on the quality of residents' discharge summaries. *J Gen Intern Med*. 2010;26(1):28–32. 10.1007/s11606-010-1465-z.

32. Fang M, Linson E, Suneja M, Kuperman EF. Impact of adding additional providers to resident workload and the resident experience on a medical consultation rotation. *BMC Med Educ*. 2017;17:44. https://doi.org/10.1186/s12909-017-0874-7.

33. Haney EM, Nicolaidis C, Hunter A, Chan BKS, Cooney TG, Bowen JL. Relationship between resident workload and self-perceived learning on inpatient medicine wards: a longitudinal study. *BMC Med Educ*. 2006;6:35. https://doi.org/10.1186/I1472-6920-6-35.

34. Ong M, Bostrom A, Vidyarthi A, McCulloch C, Auerbach A. House staff team workload and organization effects on patient outcomes in an academic general internal medicine inpatient service. *Arch Intern Med*. 2007;167(1):47–52. https://doi.org/10.1001/archinte.167.1.47.

35. Baldwin DC Jr, Daugherty SR, Ryan P. How residents view their clinical supervision: a reanalysis of classic national survey data. *J Grad Med Ed*. 2010;2(1):37–45. 10-4300/JGME-D-09-00081.1.

36. Baldwin DC Jr, Daugherty SR, Ryan PM, Yaghmour NA, Philibert I. Residents'' ratings of their clinical supervision and their self-reported medical errors: analysis of data from 2009. *J Grad Med Ed*. 2018;10(2):235–241. https://doi.org/10.4300/JGME-D-18-00200.1.

37. Singman EL, Boland MV, Tian J, Green LK, Srikumaran D. Supervision and autonomy of ophthalmology residents in the outpatient clinic in the United States II: a survey of senior residents. *BMC Med Ed*. 2019;19:202. https://doi.org/10.1186/s12909-019-1620-0.

38. Singman EL, Divya Srikumaran, Green L, Tian J, McDonnell P. Supervision and autonomy of ophthalmology residents in the outpatient clinic in the United States: a survey of ACGME-accredited programs. *BMC Med Ed*. 2017;17:105. https://doi.org/10.1186/s12909-017-0941-0.

39. ten Cate O, Chen HC Hoff RG, Peters H, Bok H, van der Schaaf M. Curriculum development for the workplace using entrustable professional activities (EPAs): AMEE guide no. 99. *Med Teach*. 2015;37(11):983–1002.

40. ten Cate O, Schwartz A, Chen HC. Assessing trainees and making entrustment decisions: on the nature and use of entrustment-supervision scales. *Acad Meds*. 2020;95(11):1662–1669. https://doi.org/10.1097/ACM.0000000000003427.

41. Stahl CC, Collins E, Jung SA, et al. Implementation of entrustable professional activities into a general surgery residency. *J Surg Ed*. 2020;77(4):739–748.

42. Shorey S, Lau TC, Lau T, Ang E. Entrustable professional activities in health care education: a scoping review. *Med Ed*. 2019;53:766–777. https://doi.org/10.1111/medu.19879.

43. George BC, Teitelbaum EN, Meyerson SL, et al. Reliability, validity and feasibility of the ZWISCH scale for the assessment of intraoperative performance. *J Surg Ed*. 2014;71(6):e90–e96.

44. Fallon WF Jr, Wears RI, Tepas JJ 3rd. Resident supervision in the operating room: does this impact on outcome. *J Trauma*. 1993;35:556–561.

45. Farnon JM, Petty LA, Georgitis E, et al. A systematic review: the effect of clinical supervision on patient

and residency education outcomes. *Acad Med*. 2012;28:428–442.

46. Finn KM, Metlay JP, Chang Y, et al. Effect of increased inpatient attending physician supervision on medical errors, patient safety, and resident education: a randomized clinical trial. *JAMA Intern Med*. 2018;178(7):952–959. https://doi.org/10.1001/jamainternmed.2018.1244.

47. Tonelli CM, Cohn T, Abdelsattar Z, Luchette FA, Baker MS. Association of resident independence with short-term clinical outcome in core general surgery procedures. *JAMA Surg*. 2023;158(3):302–309. https://doi.org/10.1001/jamasurg.2022.6971.

48. Giordano L, Oliver A, Peretti GM, Maffulli N. The presence of residents during orthopedic operation exerts no negative influence on outcome. *British Med Bulletin*. 2019;130:65–80. https://doi.org/10.1093/bmb/Idz009.

49. Christophersen C, Fonnes S, Andresen K, Rosenberg J. Risk of reoperation for recurrence after elective primary groin and ventral hernia repair by supervised residents. *JAMA Surg*. 2023;158(4):359–367. https://doi.org/10.1001/jamasurg.2022.7502.

50. Schumacher DJ, Holmboe E, Carraccio C, et al. Resident-sensitive quality measures in the pediatric emergency department: exploring relationships with supervisor entrustment and patient acuity and complexity. *Acad Med*. 2020;95(8):1256–1264. https://doi.org/10.1097/ACM.0000000000003242.

51. Kashner TM, Clarke CT, Aron DC, et al. The 9-criteria evaluation framework for perceptions survey: the case of VA's learners' perceptions survey. *Biostat Epidemiol*. 2020;4(1):140–171. https://doi.org/10.1080/24709360.2018.1553362.

APPENDIX

Definitions and Calculations of VA's Resident Workload-based Academic Performance Metrics (WRAP).

A. WRAP Metrics

A.1 Resident Clinic Workload (RCW)

RCW represents 'potential' clinical opportunities per resident. It is computed by the total clinic RVUs on a given service-date divided by the resident count and reported as RVU per resident per clinic-day.

Specifically, let W_{Tfqt} be the total RVU workload produced in clinic f on a given service-date represented by consecutively numbered service-days $t = 1, 2, \ldots, 92$ from the beginning of each annual quarter $q = 1, 2, 3, 4$, within a given academic year $T = 2016, 2017, 2018, 2019$. Let $[\mathrm{PHYR}_{Tfqt}]$ be the resident count measured in half-day shifts in clinic f on service-date (T,q,t). One resident working a half day is countered as 0.5 resident clinic-day. Then,

$$[\mathrm{RCW}]_{Tfqt} = \frac{W_{Tfqt}}{[\mathrm{PHYR}_{Tfqt}]}. \tag{1a}$$

A.2. Resident Assigned-Workload (RAW)

RAW represents clinic workload that had been assigned to a resident. It is computed by the total clinic RVUs on a given service-date that occurred during patient encounters in which a resident was listed in the patient's health record as a care provider, divided by the resident count, and reported as RVUs per resident per clinic-day.

Let W^{e}_{Tfqt} be the total assigned RVU workload in clinic f on service-date (T,q,t). Then,

$$[\mathrm{RAW}]_{Tfqt} = \frac{W^{e}_{Tfqt}}{[\mathrm{PHYR}_{Tfqt}]}. \tag{2a}$$

A.3. Resident Productivity (RP)

RP measures the productivity of residents to produce assigned workload. It is computed by the total assigned clinic RVUs that had been produced by residents in a given clinic on a given service-date, divided by the resident count, and reported as RVUs per resident per clinic-day.

Let $\widehat{W}^{e}_{Tmfqt} = g^{Tm}(\mathbf{X}_{Tfqt}|\widehat{\Gamma}_{Tmfq})$ be the expected total assigned RVU workload produced in clinic f on service-date (T,q,t) and computed from g^{Tm} as an estimated generalized linear mixed model with random coefficients indexed by academic year T and equation $m = 1, 2, \ldots, m_T$. g^{Tm} models workload by specifying all main effects, 2-way interactions, and second-order transforms of resident, attending physician, and other provider counts, by clinic and service-date - $\mathbf{X}_{Tfqt}$. $\widehat{\Gamma}_{Tmfq}$ is a vector of the respective estimated coefficients. Model specifications were selected based on a Best Approximating Model method [a] and are specified below.[b]

Let $\tilde{\mathscr{W}}^{eo}_{Tmfqt} = [\widehat{W}^{e}_{Tmfqt}] - [\widehat{W}^{e}_{Tmfqt}|\mathbf{X}_{Tfqt} = 0]$ be the expected total assigned RVUs produced by all providers and residents for facility f and service-date (T,q,t) from BAM (T,m).

Let $\tilde{\mathscr{W}}^{eR}_{Tmfqt} = [\widehat{W}^{e}_{Tmfqt}] - [\widehat{W}^{e}_{Tmfqt}|[\mathrm{PHYR}] = 0]$ be the expected [BAM]total assigned RVUs produced by residents for facility f and service-date (T,q,t) from BAM (T,m). Then,

$$[\mathrm{RP}]_{Tmfqt} = \frac{\left(\frac{\tilde{\mathscr{W}}^{eR}_{Tmfqt}}{\tilde{\mathscr{W}}^{eo}_{Tmfqt}}\right)\left[W^{e}_{Tfqt}\right]}{\left[PHYR_{Tfqt}\right]}, \tag{3}$$

$$[\mathrm{RP}]_{Tfqt} = \sum_{m \in T}\Big([\mathrm{RP}]_{Tmfqt} * z_{Tm}\Big), \tag{4}$$

where z_{TM} is how well the data fit the BAM (T,m) based on the Generalized Bayesian Information Criterion. That is, for each BAM, RP equals the ratio of expected resident-produced assigned workload to expected assigned workload times the assigned workload divided by resident count, computed per BAM, and model averaged over all BAMs weighted by the model's data fit.

A.4. Resident Net Productivity (RNP)

RNP measures resident productivity to produce assigned workload that accounts for the opportunity cost when attending physicians divert time away from clinic RVU generating activities in order to supervise residents. That is, RNP measures resident productivity net of supervision.

RNP is computed by the total assigned clinic RVUs that had been produced by residents in a given clinic on a given service-date but net of supervision, divided by the resident count, and reported as RVUs per resident per clinic-day.

Computations are similar to RP but with clinic workload W_{Tfqt} replacing clinic assigned workload W^{e}_{Tfqt}.

Let $\widehat{W}_{Tmfqt} = g^{Tm}(\mathbf{X}_{Tfqt}|\widehat{\beta}_{Tmfq})$ be the expected total clinic RVU workload produced in clinic f on service-date (T,q,t) computed from the estimated model g^{Tm} (see above).

Let $\tilde{\mathscr{W}}^{o}_{Tmfqt} = [\widehat{W}_{Tmfqt}] - [\widehat{W}_{Tmfqt}|\mathbf{X}_{Tfqt} = 0]$ be the expected total clinic RVUs produced by all providers and residents for facility f and service-date (T,Q,t) from BAM (T,m).

Let $\tilde{\mathscr{W}}^{R}_{Tmfqt} = [\widehat{W}_{Tmfqt}] - [\widehat{W}_{Tmfqt}|[\mathrm{PHYR}] = 0]$ be the expected total clinic RVUs produced by residents for facility f and service-date (T,Q,t) from BAM (T,m). Then,

$$[\mathrm{RNP}]_{Tmfqt} = \frac{\left(\frac{\tilde{\mathscr{W}}^{R}_{Tmfqt}}{\tilde{\mathscr{W}}^{O}_{Tmfqt}}\right)[W_{Tfqt}]}{[PHYR_{Tfqt}]}, \tag{5}$$

$$[\mathrm{RNP}]_{Tfqt} = \sum_{m \in T}\left([\mathrm{RNP}]_{Tmfqt} * z_{Tm}\right). \tag{6}$$

A.5. Attending Physician Productivity (PP)

PP measures the productivity of attending physicians on a given service-date. It is calculated by the total clinic RVUs attending physicians produced on a given service-date, divided by the attending physician count, and reported as RVU per physician per clinic-day.

Let $\tilde{\mathscr{W}}^{A}_{Tmfqt} = [W_{Tfqt}] - [\hat{W}_{Tfqt}|[PHY] = 0\ ]$ be the expected attending physicians total contribution to clinic RVUs computed from the estimated total clinic workload production model. Then,

$$[\mathrm{PP}]_{Tmfqt} = \frac{\left(\frac{\tilde{\mathscr{W}}^{A}_{Tmfqt}}{\tilde{\mathscr{W}}^{O}_{Tmfqt}}\right)[W_{Tfqt}]}{[PHY_{Tfqt}]}, \tag{7}$$

$$[\mathrm{PP}]_{Tfqt} = \sum_{m \in T}\left([\mathrm{PP}]_{Tmfqt} * z_{Tm}\right). \tag{8}$$

A.6. Average Nonsupervising Physician Productivity (avgNSPP)

The attending physician productivity averaged over all clinics and service-dates when residents were not present.

$$[avgNS\mathrm{PP}] = \mathrm{avg}\left([\mathrm{PP}]_{Tfqt}|[\mathrm{PHYR}] = 0\right). \tag{9}$$

B. WRAP Rates

B.1. Workload Assigned Rate (WArate)

The Workload Assigned Rate is the first component to the Workload Index that benchmarks resident assigned workload to the productivity of the average attending physician producing care with no resident supervising responsibilities. It is computed by the Resident Assigned-Workload per resident clinic-day (RAW) divided by the non-supervising average physician productivity (*avgNS*PP), and reported as a percent or proportion.

$$[\mathrm{WArate}]_{Tfqt} = \frac{[\mathrm{RAW}]_{Tfqt}}{[avgNS\mathrm{PP}]}. \tag{10}$$

B.2. Workload Produced Rate (WProrate)

The Workload Produced Rate is the second component to the Workload Index that measures the extent residents produced all of their assigned RVU workload. It is computed by the total assigned clinic RVUs that residents produced per resident clinic-day (RP) divided by the Resident Assigned-Workload per resident clinic-day (RAW) and reported as a percent or proportion.

$$[\mathrm{WProrate}]_{Tmfqt} = \frac{[\mathrm{RP}]_{Tmfqt}}{[\ \mathrm{RAW}]_{Tfqt}}, \tag{11}$$

$$[\mathrm{WProrate}]_{Tfqt} = \sum_{m \in T}\left([\mathrm{WProrate}]_{Tmfqt} \times z_{Tm}\right). \tag{12}$$

B.3. Physician Productivity Rate (PProrate)

The Physician Productivity Rate benchmarks attending physician productivity to the productivity of the average attending physician providing patient care with no resident supervising responsibilities. It is computed by the physician productivity metric (PP) divided by the non-supervising average attending physician productivity (*avgNS*PP) and reported as a percent or proportion.

$$[\mathrm{PProrate}]_{Tfqt} = \frac{[PP]_{Tfqt}}{[avgNS\mathrm{PP}]}. \tag{13}$$

C. WRAP Indices

C.1. Supervision Index (SI)

The Supervision Index (progressive autonomy) measures the extent residents produced assigned clinic RVUs that was progressively autonomous from direct oversight of attending physicians. It is calculated by Resident Net Productivity (RNP) divided by Resident Productivity (RP) and reported as a percent or proportion.

$$[\mathrm{SI}]_{Tmfqt} = \frac{[\mathrm{RNP}]_{Tmfqt}}{[\mathrm{RP}]_{Tfqt}}, \tag{14}$$

$$[\mathrm{SI}]_{Tfqt} = \sum_{m \in T}\left([\mathrm{SI}]_{Tmfqt} \times z_{Tm}\right). \tag{15}$$

C.2. Workload Index (WI)

The Workload Index measures resident productivity as a percent of the productivity of the average attending physician on non-teaching clinic-days. It is computed by resident productivity (RP) divided by the nonsupervising average attending physician productivity (*avgNS*PP) and reported as a percent or proportion. It is also the product of workload assigned and produced rates.

$$[\mathrm{WI}]_{Tfqt} = \frac{[\mathrm{RP}]}{[avgNS\mathrm{PP}]} = \frac{[\mathrm{RAW}]}{[avgNS\mathrm{PP}]} \times \frac{[\mathrm{RP}]}{[\mathrm{RAW}]}$$
$$= [\mathrm{WArate}] \times [\mathrm{WPrate}]. \quad (16)$$

C.3. Readiness Index (to Enter Independent Practice) (RI)

The Readiness Index measures the extent resident productivity net of supervision approaches the productivity of the average attending physician on clinic-days when residents are absent. It is calculated by resident net productivity (RNP) divided by the attending physician productivity averaged over non-teaching clinic service-dates (*avgNS*PP) and reported as a percent or proportion. It combines workload and supervision indices.

$$[\mathrm{RI}]_{Tfqt} = \frac{[\mathrm{RNP}]_{Tfqt}}{[avgNS\mathrm{PP}]} = \frac{[\mathrm{RAW}] \times \frac{[\mathrm{RP}]}{[\mathrm{RAW}]} \times \frac{[\mathrm{RNP}]}{[\mathrm{RP}]}}{[avgNS\mathrm{PP}]}$$
$$= [\mathrm{WI}] \times [\mathrm{SI}]. \quad (17)$$

List of Abbreviations

*avgNS*PP	Non-Supervising average Physician Productivity
BAM	Best Approximating Model
CPT	Current Procedure Terminology
GBIC	Generalized Bayesian Information Criterion
GME	Graduate Medical Education
HEDIS	Healthcare Effectiveness Data and Information Set
ICC	Intraclass Correlation Coefficient
ID	Interdecile range
PGY	Post Graduate Year
PP	Physician Productivity
PPrate	Physician Productivity Rate
RAW	Resident Assigned-Workload
RI	Readiness Index
RCW	Resident Clinic-Workload
RNP	Resident Net Productivity
RP	Resident Productivity
RVU	Relative Value Unit
SD	standard deviation
VA	Department of Veterans Affairs
WArate	Workload Assigned Rate
WPrate	Workload Produced Rate
WI	Workload Index
WRAP	Workload-based Resident Academic Performance Indices and Metrics

a– The generalized linear mixed model g^{Tm} has a different specification m = 1, 2, …, mT for each academic year T. These specifications were selected based on the Best Approximating Model method (BAM). BAM identifies the best fitting models from a search of all possible models that are linear in coefficients. The procedure searches exhaustively over all resident-provider counts as main effects, plus statistically significant 2-way interactions and second order terms. Significant (p = .05) interactions from among the candidate interaction pool were identified using the robust sandwich estimator where the constituent main effects for the interaction were present in each individual model. The m_T best fitting models based on the Generalized Bayesian Information Criterion (GBIC) that were observationally equivalent at the p = .05 level were selected to form an "Occam's Window" for each respective academic year. Exhaustively searched model space size by academic year was 131,072 models for 2016, 131,072 models for 2017, 132,768 models for 2018, and 132,768 models for 2019.

b– Models are described here. Symbols are defined in Table 2. All models include (PHY), (PHY-R), (PHY)*(PHY-R). Additional variables are:

T = 2019, m = 1: (RN), (NP), (OTH), (NP)*(OTH);

T = 2019, m = 2: (RN), (NP), (OTH), (RN)*(PHY-R), (NP)*(OTH);

T = 2018, m = 1: (RN), (OTH), (PA)*(PHY-R), (RN)*(OTH);

T = 2018, m = 2: (RN), (NP), (OTH), (PA)*(PHY-R), (RN)*(OTH);

T = 2018, m = 3: (RN), (OTH), (RN)*(OTH);

T = 2018,m = 4: (RN), (NP), (OTH), (RN)*(OTH);

T = 2018, m = 5: (PA), (RN), (OTH), (RN)*(OTH);

T = 2018, m = 6: (RN), (NP), (OTH), (PA)*(NP), (PA)*(PHY-R), (RN)*(OTH);

T = 2018, m = 7: (RN), (OTH), (RN)*(OTH), (RN)*(PHY);

T = 2017, m = 1: (RN), (OTH), (PA)*(RN), (PA)*(PHY), (RN)*(OTH), (NP)*(OTH), (OTH)*(PHY);

T = 2017, m = 2: (RN), (OTH), (PA)*(OTH), (PA)*(PHY), (RN)*(OTH), (NP)*(OTH), (OTH)*(PHY);

T = 2017, m = 3: (RN), (OTH), (PA)*(PHY), (RN)*(OTH), (NP)*(OTH), (OTH)*(PHY);

T = 2017, m = 4: (RN), (OTH), (PA)*(PHY), (RN)*(OTH), (NP)*(OTH);

T = 2017, m = 5: (PA), (RN), (OTH), (PA)*(RN), (PA)*(PHY), (RN)*(OTH), (NP)*(OTH), (OTH)*(PHY);

T = 2017, m=6: (PA), (PA)*(PHY), (RN), (RN)*(OTH), (NP)*(OTH), (OTH), (OTH)*(PHY);

T = 2017, m = 7: (PA)*(RN), (PA)*(OTH), (PA)*(PHY), (RN), (RN)*(OTH), (NP)*(OTH), (OTH), (OTH)*(PHY);

T = 2017, m = 8: (RN), (OTH), (PA)*(RN), (PA)*(PHY), (RN)*(OTH), (NP)*(OTH);

T = 2016, m = 1: (PA), (RN), (OTH), (PA)*(RN), (PA)*(OTH), (PA)*(PHY-R), (NP)*(OTH), (NP)*(PHY);

T = 2016, m = 2: (RN), (OTH), (PA)*(RN), (PA)*(PHY-R), (NP)*(OTH), (NP)*(PHY);

T = 2016, m = 3: (RN), (OTH), (PA)*(RN), (PA)*(OTH), (PA)*(PHY-R), (NP)*(OTH), (NP)*(PHY);

T = 2016, m = 4: (PA), (RN), (OTH), (PA)*(RN), (PA)*(OTH), (PA)*(PHY-R), (NP)*(OTH), (NP)*(PHY), (OTH)*(PHY).