

2016

Benchmarking Academic Anatomic Pathologists: The Association of Pathology Chairs Survey

Barbara S. Ducatman MD
West Virginia University

Tristram Parslow MD, PhD
Emory University

Follow this and additional works at: https://researchrepository.wvu.edu/faculty_publications



Part of the [Medicine and Health Sciences Commons](#)

Digital Commons Citation

Ducatman, Barbara S. MD and Parslow, Tristram MD, PhD, "Benchmarking Academic Anatomic Pathologists: The Association of Pathology Chairs Survey" (2016). *Faculty & Staff Scholarship*. 2462.
https://researchrepository.wvu.edu/faculty_publications/2462

This Article is brought to you for free and open access by The Research Repository @ WVU. It has been accepted for inclusion in Faculty & Staff Scholarship by an authorized administrator of The Research Repository @ WVU. For more information, please contact ian.harmon@mail.wvu.edu.

Benchmarking Academic Anatomic Pathologists: The Association of Pathology Chairs Survey

Academic Pathology
 Volume 3: 1–12
 © The Author(s) 2016
 Reprints and permission:
sagepub.com/journalsPermissions.nav
 DOI: 10.1177/2374289516666832
apc.sagepub.com



Barbara S. Ducatman, MD¹, and Tristram Parslow, MD, PhD²

Abstract

The most common benchmarks for faculty productivity are derived from Medical Group Management Association (MGMA) or Vizient-AAMC Faculty Practice Solutions Center[®] (FPSC) databases. The Association of Pathology Chairs has also collected similar survey data for several years. We examined the Association of Pathology Chairs annual faculty productivity data and compared it with MGMA and FPSC data to understand the value, inherent flaws, and limitations of benchmarking data. We hypothesized that the variability in calculated faculty productivity is due to the type of practice model and clinical effort allocation. Data from the Association of Pathology Chairs survey on 629 surgical pathologists and/or anatomic pathologists from 51 programs were analyzed. From review of service assignments, we were able to assign each pathologist to a specific practice model: general anatomic pathologists/surgical pathologists, 1 or more subspecialties, or a hybrid of the 2 models. There were statistically significant differences among academic ranks and practice types. When we analyzed our data using each organization's methods, the median results for the anatomic pathologists/surgical pathologists general practice model compared to MGMA and FPSC results for anatomic and/or surgical pathology were quite close. Both MGMA and FPSC data exclude a significant proportion of academic pathologists with clinical duties. We used the more inclusive FPSC definition of clinical "full-time faculty" (0.60 clinical full-time equivalent and above). The correlation between clinical full-time equivalent effort allocation, annual days on service, and annual work relative value unit productivity was poor. This study demonstrates that effort allocations are variable across academic departments of pathology and do not correlate well with either work relative value unit effort or reported days on service. Although the Association of Pathology Chairs–reported median work relative value unit productivity approximated MGMA and FPSC benchmark data, we conclude that more rigorous standardization of academic faculty effort assignment will be needed to improve the value of work relative value unit measurements of faculty productivity.

Keywords

anatomic pathology, benchmarking, clinical effort, Medical Group Management Association, productivity, surgical pathology, Vizient-AAMC Faculty Practice Solutions Center, work relative value units

Received August 05, 2016. Received revised August 08, 2016. Accepted for publication August 08, 2016.

Introduction

With the changing financial environment of health care, academic medical centers face intense pressure to maximize productivity and efficiency and to constrain costs. This in turn requires assessment and accountability of the productivity of departments and individual faculty for their clinical productivity. Faced with the challenge of quantifying surgical, primary-care, psychiatric, and other disparate professional activities across specialties and institutions, academic leaders and their

¹ Department of Pathology, West Virginia University School of Medicine, Morgantown, WV, USA

² Department of Pathology, Emory University School of Medicine, Atlanta, GA, USA

Corresponding Author:

Barbara S. Ducatman, Department of Pathology, West Virginia University School of Medicine, PO Box 9203, Morgantown, WV 26506, USA.
 Email: bducatman@hsc.wvu.edu



financial managers increasingly rely on a metric that purports to encapsulate the training and effort required for a physician to deliver clinical service across disciplines. The physician work relative value unit (wRVU), originally conceived by Hsiao et al for standardizing workloads and payment for physician services across specialties, is now central to the formulas that determine Medicare payments for those activities, with indirect effects on reimbursements by other payers.^{1,2} Specific wRVU values have been assigned to each of the clinical activities defined under the common procedural terminology (CPT) classification scheme.³

With the recognition that reimbursement per se is an imperfect measure of a given physician's clinical contributions owing to variability in contractual revenue recovery, wRVUs are used to assess productivity. Budgets, faculty performance evaluations, and incentive plans of physician practices now commonly focus on wRVU production and often include targets that are set in reference to wRVU benchmarks, usually from Medical Group Management Association (MGMA) and Vizient-AAMC Faculty Practice Solutions Center[®] (FPSC).¹

Since these are the accepted methods of benchmarking physician workload, it is critical that wRVU benchmarks be valid, standardized, reproducible, and generalizable. We obtained summary data from MGMA and FPSC to compare with our analyses. The MGMA collects data on both community and academic pathologists, and FPSC collects data only on academic pathologists. However, although we compared our data to these benchmarks, we are not permitted to publish the benchmarks (these are proprietary data).

We hypothesized that data from the Association of Pathology Chairs (APC) survey could be used to validate and improve the MGMA and FPSC data sets for academic pathology faculty productivity. In order to do this, we analyzed our data using the methods by which MGMA and FPSC calculate their median benchmarks. In addition, we reviewed the interaction between clinical effort allocation (clinical full-time equivalent [cFTE]), days on service (a marker of effort allocation), and wRVU data.

Materials and Methods

The APC Survey Design

The APC survey is issued annually by the APC practice and management committee and all member departments of the APC are invited to complete it. Working closely with chairs and departmental administrators who share a nuanced understanding of the specialty, the APC has refined its survey through successive iterations, with the goals of increasing participation, standardizing terminology and methodology, and eliminating common sources of error. The survey is usually filled out by the departmental administrator based on retrospective data from each department's most recently completed fiscal year and is then returned as an electronic spreadsheet to the APC, where institutions are de-identified and data compiled. Only the 2 most recent (2013 and 2014) surveys requested

anonymized wRVU, effort allocation, academic rank, days on service, and other professional data on individual faculty. Whenever possible, the data on wRVUs, effort allocation, and days on service were split by subspecialty for those pathologists providing clinical service in more than 1 subspecialty area.

The APC Survey Data Analysis

For the 2014 survey, one of us (B.S.D.) contacted the administrators of departments that had submitted evident outliers to correct potential errors and ensure that uniform methodology had been applied; data for fewer than 5% of the 2014 subjects were corrected as a result. Since the data are completely anonymized and de-identified as to program and individual, it was exempt from institutional review board approval. The present study used data from the 2014 survey only, which was merged into a single Excel spreadsheet.^a

Our analysis includes all pathologists who were designated to practice "anatomic" (AP), "surgical" pathology (SP), and/or a subspecialty or a "hybrid" combination of AP subspecialties but excluded those who practice hematopathology, autopsy, or forensic pathology exclusively. Departments self-designated their pathologists' practices without guidelines. We assume there is considerable heterogeneity in departmental individual practices and this will affect the overall derived benchmarks. Thus, some departments may have lumped faculty who practice SP and cytopathology as "surgical/anatomic pathologists" for effort allocation and wRVUs, while others may have split these 2 groups. However, the same issues apply to the MGMA and FPSC data sets.

Comparison With MGMA and FPSC Data and Analyses

We received summary data and an understanding of the methodology of how MGMA calculates their benchmarks from a senior medical school administrator for West Virginia University, who has used the data for many years. As part of our effort to improve benchmarking in pathology, one of us (B.S.D.) traveled to Chicago in December 2015 and discussed the data and methodology with the staff at FPSC. During this visit, we clarified the FPSC definitions for AP and SP (self-described) and the methodology.

Both MGMA and FPSC have used reasonable but arbitrary definitions for a cFTE: MGMA uses a cutoff value of 0.67 cFTE and then does a distribution analysis of actual wRVUs. In contrast, FPSC uses a cutoff value of 0.6 cFTE and further normalizes the data (dividing wRVUs by cFTE). In 2015, FPSC will use a threshold cFTE of 0.5. We present APC benchmarks using both methods. In this fashion, we sought to gain insight into the value and limitations of the benchmark data provided by these organizations.

Definition of Practice Types in AP

We divided the entire set of those practicing AP into 4 groups by practice model:

1. General practice model includes pathologists only practicing general SP/AP (as defined by the program administrator).
2. Hybrid includes pathologists practicing general SP/AP plus a subspecialty.
3. Single subspecialty includes pathologists practicing only 1 AP subspecialty.
4. Multiple subspecialties includes pathologists practicing more than 1 anatomic subspecialty.

Statistical Analysis

Data were imported to the JMP statistical package^b for analysis using nonparametric techniques (since the data did not follow a normal or parametric distribution). Thus, the most common statistical test used was the Wilcoxon Kruskal-Wallis test for nonparametric data by pairs.

In our presented data, the sum of all effort allocations (ie, part A, part B, education, research, and service) at the 25th, 50th, and 75th percentile do not add up to 1.0 full-time equivalent (FTE). This would be expected in independent distributions, since some individuals do not have any effort dedicated to educational, research, and/or service activities, and we only included faculty members with effort allocation greater than 0 in our distribution analyses.

Results

Overall Study Population

Anonymized data for 2014 were received on a total of 1280 individual faculty from 53 North American pathology departments that responded to the APC survey. Departments had been asked to designate 1 or more categories of professional activity for each faculty member from a list of options (ie, including but not limited to part A, part B, education, research, and nonclinical service) as well as days assigned to clinical service and salary data. The present study focused only on 629 faculty members from 51 programs who had been identified as devoting any or all of their clinical effort allocation (part B effort or cFTE) to clinical service in AP in any practice model and had earned any wRVUs; many of these faculty were also designated as providing part A services, and most were also engaged in teaching, research, administrative, or other duties. Thus, we excluded 584 faculty pathologists from further analysis since they practiced only clinical pathology or autopsies or did not report any part B effort or wRVUs.

Entire AP Data Set

The 629 faculty under study were reported to devote a median of 0.7 cFTE professional effort to part B services (range: 0.05-

1.0 cFTE), to provide a median of 148 days of on-service duty (range: 5-365 days) to such services, and to generate a global median of 4544 wRVUs (over the total range of effort allocation with a range of 4-14 160 wRVUs) in 2014. Many of these individuals also had teaching, research, and/or administrative time of greater than 0.10 FTE in aggregate.

Of the 553 anatomic pathologists whose gender was reported, 269 (49%) were women and 284 were men (51%). Data on academic rank were provided for 614 individuals, comprising 5 instructors (<1%), 252 assistant professors (43%), 166 associate professors (28%), 189 professors (29%), and 2 emeritus faculty (<1%). Data on type of practice, median effort allocation for part A, part B, education and research, as well as assigned days on service, wRVU generation, and salaries for this group, stratified by rank, are summarized in Table 1. Among these parameters, practice models and median salaries varied significantly among all 3 ranks. Median part B effort allocation was significantly higher for assistant and associate professors as compared to full professors, and median wRVUs were significantly lower for full professors as compared to assistant and associate professors. In addition, days on service varied significantly only between assistant and full professors. Part A effort, educational effort, and research effort were not significantly different across the 3 ranks.

Effects of Practice Model

There were significant differences between the 4 practice models (see Table 2). The majority of academic anatomic pathologists are fairly evenly distributed across the general practice (27%), hybrid (31%), and single subspecialty (28%) models, while fewer practice in the multiple subspecialties model (14%). Overall, 72% of academic anatomic pathologists spend at least some of their time practicing a subspecialty. However, 58% of academic pathologists still spend at least some of their time signing out general SP. Whether these practice patterns result from the increasing trend of residents pursuing multiple fellowships or are the cause cannot be determined from this data. In this sample group, the most common subspecialties practiced as a single subspecialty were neuropathology (n = 43), cytopathology (n = 31), renal (n = 23), pediatric pathology (n = 21), dermatopathology (n = 18), and gastrointestinal (n = 10). The remaining single subspecialties had fewer than 10 pathologists in each category.

The largest difference in the median part B cFTE was for single subspecialty practice (0.55 FTE), which was significantly lower than that for the general practice, hybrid practice, or multiple subspecialties models. Differences between the other practice models were not statistically significant. In contrast, median days on service were significantly different among all practice models except between the hybrid and single subspecialty practice models. Median values for wRVUs were highest among the general practice and multiple subspecialties practice models, with significantly lower median wRVUs for the hybrid and single subspecialty models. Finally median research effort was lowest in the multiple subspecialties model and significantly higher in the general and single

Table 1. Differences Between Academic Ranks.

	Assistant, N = 252, n (%)	Associate, N = 166, n (%)	Full, N = 189, n (%)	P Value
Type of practice by rank				
General practice	71 (44%)	53 (31%)	43 (25%)	<.004*
Hybrid practice	91 (48%)	50 (26%)	48 (24%)	*
Single subspecialty practice	50 (30%)	41 (25%)	73 (45%)	*
Multiple subspecialties practice	40 (46%)	22 (25%)	25 (29%)	*
Median data				
Part A effort (if > 0) with number (%) of that rank with any effort	0.13, n = 95 (38%)	0.15, n = 91 (55%)	0.2, n = 93 (49%)	NS [§]
Part B effort FTE	0.7 [†]	0.7 [^]	0.53 ^{†^}	<.0001 ^{†^}
Days on service	158 [†]	150	130 [†]	<.005 [†]
wRVUs	5076 [†]	4706 [^]	3394 ^{†^}	<.0001 ^{†^}
Educational effort	0.13, n = 144 (57%)	0.11, n = 109 (66%)	0.1, n = 129 (68%)	NS [§]
Research effort	0.11, n = 97 (38%)	0.1, n = 74 (45%)	0.1, n = 91 (48%)	NS [§]
Total salary	US\$200 391	US\$243 176	US\$289 405	<.0001 [†] for each pair

Abbreviations: FTE, full-time equivalent; wRVUs, work relative value units.

*Statistically significant by chi-square (applies for all 4 rows).

[§]Not statistically significant for any pair by the Wilcoxon Kruskal-Wallis test for nonparametric data by pairs.

^{†^}Statistically significant by the Wilcoxon Kruskal-Wallis test for nonparametric data by pairs at the P value shown.

Table 2. Differences Between Practice Models.

Median Values	General n = 170 (27%)	Hybrid n = 193 (30%)	Single Subspecialty n = 176 (28%)	Multiple Subspecialties n = 90 (14%)	P Value
Part A effort (FTE) if > 0 (percentage > 0)	0.11 (50%)	0.19 (46%)	0.19 (46%)	0.15 (48%)	NS*
Part B effort (cFTE)	0.7 [†]	0.7 [‡]	0.55 ^{†‡§}	0.67 [§]	<.003 [†] , <.0001 [‡] , <.006 [§]
Days on service	130 ^{†‡§}	152 [‡]	150 ^{§^}	171 ^{†^}	<.0001 [†] , <.0003 [‡] , <.02 [§] , <.01 [^]
wRVUs	5215 ^{‡#}	4479 ^{†‡^}	2662 ^{§^#}	5426 ^{†§}	<.001 ^{†‡} , <.0001 ^{§^#}
Educational effort if > 0 (percentage > 0)	0.12 (56%)	0.1 (67%)	0.12 (66%)	0.12 (61%)	NS*
Research effort if > 0 (percentage > 0)	0.15 [^] (44%)	0.1 (42%)	0.15 [†] (44%)	0.05 ^{†^} (43%)	<.0006 [†] , <.004 [^]
Median total salary	\$221 216	\$234 171	\$241 500	\$235 000	NS*

Abbreviations: FTE, full-time equivalent; cFTE, clinical full-time equivalent; wRVUs, work relative value units.

*Not significant for any pair by Wilcoxon Kruskal-Wallis test for nonparametric data by pairs.

^{†‡§#}Statistically significant by Wilcoxon Kruskal-Wallis test for nonparametric data by pairs at the P value shown.

subspecialty models. There were no statistically significant differences in part A median effort, educational effort, and salary among the practice models.

The wRVU Data

Figure 1A depicts a 2-dimensional scatterplot with annual wRVUs plotted against the reported part B service effort for all general practice faculty members with reported effort and wRVUs above 0. The nonparametric density plot superimposed on this scattergram has the darkest contour showing the most dense concentration of data points (top quartile), while the lightest contour defines the least concentrated points (first quartile). The line of regression and its equation

and R^2 value are also shown in the figure. The faculty pathologists that would be considered “full-time” by FPSC (≥ 0.6 cFTE) are enclosed in a rectangle. Figure 1B shows a second scattergram composed only of these FPSC model full-time pathologists. Not surprisingly, the graph is similar for all effort allocations; however, the line of regression is not as steep (or for those pathologists there is a lesser difference between wRVUs, that is, from 0.6 to 1.0 cFTE than the difference from 0.05 to 1.0 cFTE). The model is also less predictive as seen in the R^2 results. The dense cluster seen in Figure 1A is more prominent in Figure 1B and there is less variability. The densest cluster contains those with cFTE between 0.65 and 0.75 and wRVU between approximately 4200 and 7000. Compared to other practice models, the

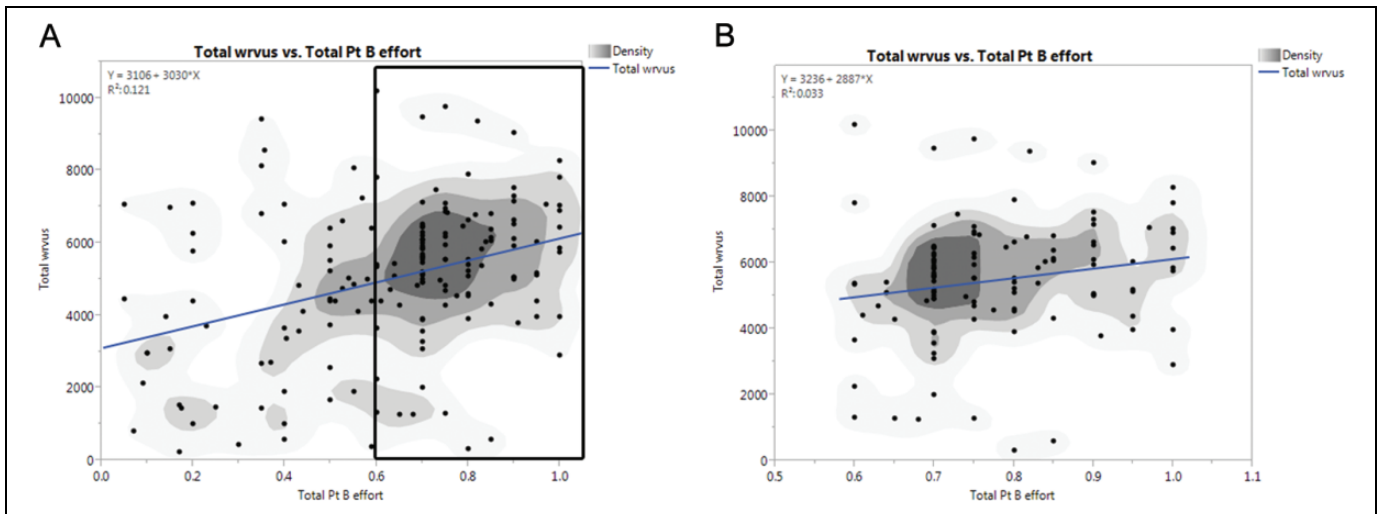


Figure 1. A, Scattergram of part B effort versus wRVUs with a nonparametric density analysis for faculty pathologists across all clinical effort allocations in the general practice model. In the rectangular box are included those pathologists who would be in the FPSC (≥ 0.6 cFTE) analysis, which would also include the MGMA analysis (≥ 0.67 cFTE). B, Scattergram of part B effort versus wRVUs with a nonparametric density analysis for faculty pathologists in the general practice model, who would be included in the FPSC analysis (≥ 0.6 cFTE). The darkest color (top quartile) shows the densest cluster (top quartile) of points, each of which represents a faculty pathologist, while the lightest shows the least concentrated (bottom quartile). The intermediate colors show the second and third quartiles. The line of regression with its associated equation and coefficient of determination (R^2) are noted in the upper left corner. FPSC indicates Vizient-AAMC Faculty Practice Solutions Center; cFTE indicates clinical full-time equivalent; MGMA, Medical Group Management Association; wRVUs, work relative value units.

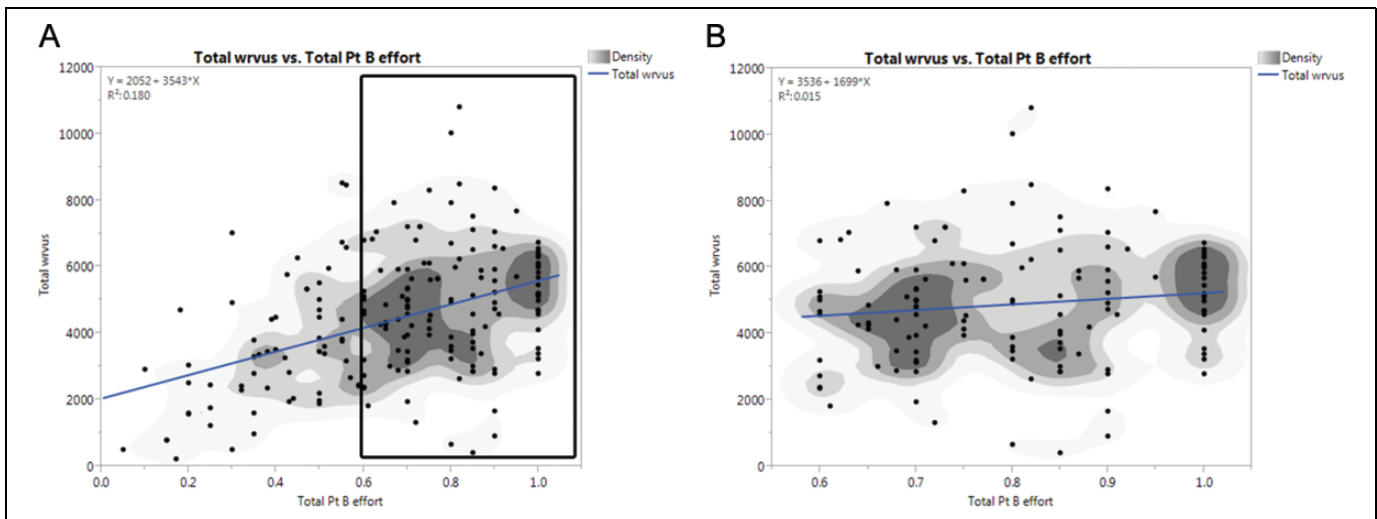


Figure 2. A, Scattergram of part B effort versus wRVUs with a nonparametric density analysis for faculty pathologists across all clinical effort allocations in the hybrid practice model. In the rectangular box are included those pathologists who would be in the FPSC (≥ 0.6 cFTE) analysis which would also include the MGMA analysis (≥ 0.67 cFTE). B, Scattergram of part B effort versus wRVUs with a nonparametric density analysis for faculty pathologists in the hybrid practice model who would be included in the FPSC analysis (≥ 0.6 cFTE). The darkest color (top quartile) shows the densest cluster (top quartile) of points, each of which represents a faculty pathologist, while the lightest shows the least concentrated (bottom quartile). The intermediate colors show the second and third quartiles. The line of regression with its associated equation and coefficient of determination (R^2) are noted in the upper left corner. FPSC indicates Vizient-AAMC Faculty Practice Solutions Center; cFTE indicates clinical full-time equivalent; MGMA, Medical Group Management Association; wRVUs, work relative value units.

general practice model is the most homogeneous and thus is most amenable to benchmarking.

We also completed similar analyses for the hybrid (Figure 2A and B), single subspecialty (Figure 3A and B), and multiple subspecialty (Figure 4A and B) practice models.

Analysis of these graphs shows the difficulty in analyzing all anatomic pathologists as a single group. The hybrid model is not clustered uniformly and demonstrates more variability than the general practice model. The single subspecialty model demonstrates multiple dense clusters, including a

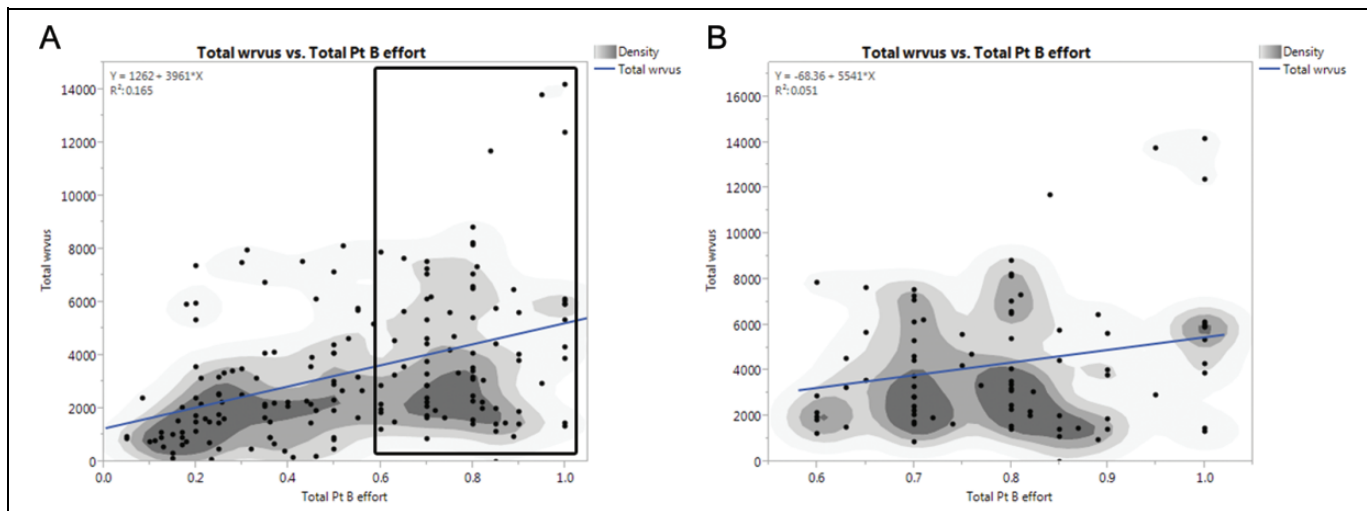


Figure 3. A, Scattergram of part B effort versus wRVUs with a nonparametric density analysis for faculty pathologists across all clinical effort allocations in the single subspecialty practice model. In the rectangular box are included those pathologists who would be in the FPSC (≥ 0.6 cFTE) analysis, which would also include the MGMA analysis (≥ 0.67 cFTE). B, Scattergram of part B effort versus wRVUs with a nonparametric density analysis for faculty pathologists in the single subspecialty practice model, who would be included in the FPSC analysis (≥ 0.6 cFTE). The darkest color (top quartile) shows the densest cluster (top quartile) of points, each of which represents a faculty pathologist, while the lightest shows the least concentrated (bottom quartile). The intermediate colors show the second and third quartiles. The line of regression with its associated equation and coefficient of determination (R^2) are noted in the upper left corner. FPSC indicates Vizient-AAMC Faculty Practice Solutions Center; cFTE indicates clinical full-time equivalent; MGMA, Medical Group Management Association; wRVUs, work relative value units.

significant number of pathologists practicing at or below 0.5 cFTE and lower wRVUs overall. The multiple subspecialties model is quite different from the single subspecialty model and more resembles the general and hybrid practice patterns. We postulate that the differences in the latter 3 groups are due to the differences in patterns for various subspecialties.

Comparison to MGMA and FPSC Data

Table 3 shows distribution analyses for each practice model including all clinical faculty (cFTE > 0 and wRVUs > 0) which were lower than either FPSC or MGMA, since these data sets exclude those pathologists of less than 0.6 cFTE and 0.67 cFTE, respectively. We also used MGMA and FPSC methodology to calculate these values in our data using their cutoffs. Our median wRVU result for the general practice model using MGMA methods was very similar (with a 2% difference) to the MGMA benchmark for AP (which as proprietary data we cannot publish). However, this included only 58% of general anatomic/surgical pathologists who practice ≥ 0.67 cFTE. In contrast, 61% of full-time (≥ 0.60 FTE) pathologists in our APC cohort were included in the FPSC-type analysis (shown in blue in the table). There was also close concordance (with a 3% difference) between the APC median for general practice with the FPSC median for AP and a somewhat greater difference (7%) with the FPSC median for SP (these are self-defined by programs in the FPSC data set). There are other differences in FPSC's data collection and analysis that likely account for the slightly larger differences. Note that the FPSC method results were higher than MGMA; this is due to the effect of normalizing wRVUs

for effort allocation. This table also shows both the effect of excluding a subset of our pathology faculty and the added effect of using normalized wRVUs on productivity benchmarks and how these analyses exclude considerable numbers of pathologists from analysis.

Clinical Effort Allocation Versus Days on Service

Since we hypothesized that an assigned effort allocation might vary among departments, we decided to compare it with the assigned days on service for full-time pathologists, again using the more inclusive FPSC benchmark for full-time faculty. Time on service should represent a discrete and clear marker for clinical effort for all 4 practice types. When we ran the general practice model, we found an outlier with a large cluster of pathologists ($n = 23$) from 1 program, all of whom practiced SP only, had 40 days of service, a cFTE of 0.7, and wRVUs between 1966 and 7113. Since the days on service did not match the cFTE and wRVUs, we excluded this program from further analysis for time on service. With this exclusion, the scattergram showed the densest cluster of general anatomic pathologist practiced between 125 and 200 days with a cFTE between 0.65 and 0.8 (Figure 5A). Furthermore, there appeared to be little if any correlation between days on service and cFTE. The line of regression was around 150 days (with a slightly negative slope). We then analyzed wRVUs against days on service for this group (Figure 5B). Here the densest cluster was between 125 to 175 days and wRVUs between about 5000 and 7000. This suggests that days on service might be useful as a surrogate for cFTE in the general practice model.

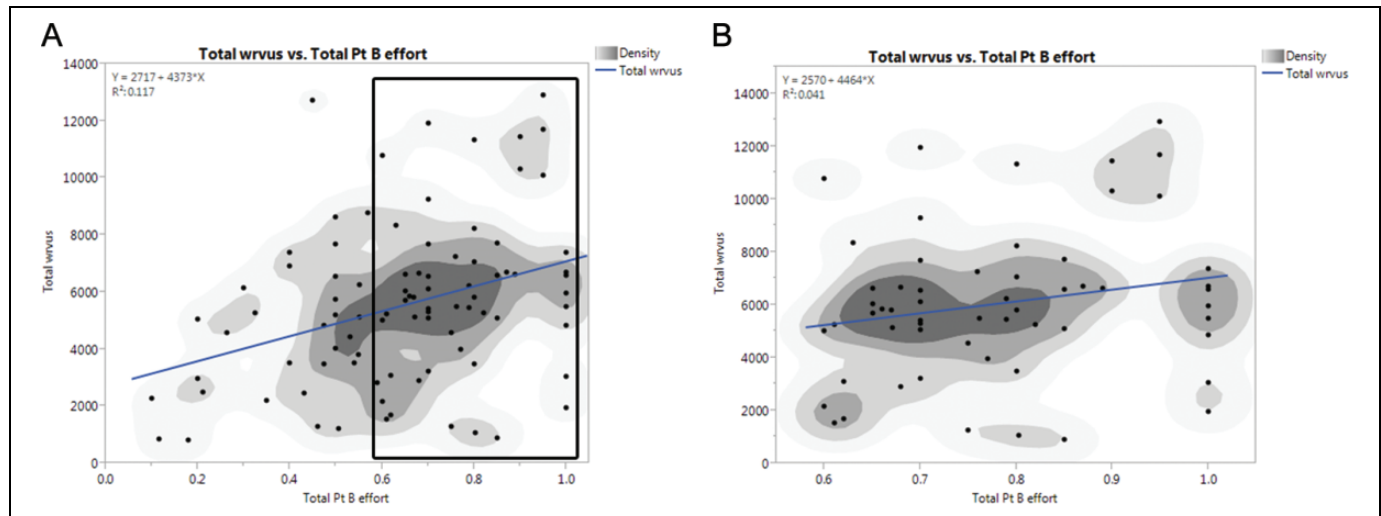


Figure 4. A, Scattergram of part B effort versus wRVUs with a nonparametric density analysis for faculty pathologists across all clinical effort allocations in the multiple subspecialties practice model. In the rectangular box are included those pathologists who would be in the FPSC (≥ 0.6 cFTE) analysis, which would also include the MGMA analysis (≥ 0.67 cFTE). B, Scattergram of part B effort versus wRVUs with a nonparametric density analysis for faculty pathologists in the multiple subspecialties practice model, who would be included in the FPSC analysis (≥ 0.6 cFTE). The darkest color (top quartile) shows the densest cluster (top quartile) of points, each of which represents a faculty pathologist, while the lightest shows the least concentrated (bottom quartile). The intermediate colors show the second and third quartiles. The line of regression with its associated equation and coefficient of determination (R^2) are noted in the upper left corner. FPSC indicates Vizient-AAMC Faculty Practice Solutions Center; cFTE indicates clinical full-time equivalent; MGMA, Medical Group Management Association; wRVUs, work relative value units.

Table 3. The APC Benchmarks With the Percentage of Faculty Included by Methodology.*,†,‡

	N (%)	Service Days (median)	25th	50th	75th
APC data: general practice model					
MGMA method	98 (58%)	144	4892	5786	6513
FPSC method	111 (65%)	144	5742	7220	8359
wRVUs for all effort allocations: actual	170 (100%)	130	3896	5215	6403
APC data: hybrid model					
MGMA method	112 (58%)	162	3645	5014	6086
FPSC method	132 (70%)	162	4565	6056	7601
wRVUs for all effort allocations: actual	193 (100%)	152	3081	4479	5746
APC data: single subspecialty model					
MGMA method	64 (36%)	180	2022	3837	6085
FPSC method	84 (48%)	177	2733	4593	7401
wRVUs for all effort allocations: actual	176 (100%)	150	1499	2662	5046
APC data: multiple subspecialties model					
MGMA method	44 (49%)	200	5012	6023	7460
FPSC method	58 (64%)	191	5408	7663	9584
wRVUs for all effort allocations: actual	90 (100%)	171	3470	5426	6752

Abbreviations: APC, Association of Pathology Chairs; cFTE, clinical full-time equivalent; MGMA, Medical Group Management Association; FPSC, Vizient-AAMC Faculty Practice Solutions Center; wRVUs, work relative value units.

*MGMA method in red: data included only for cFTE ≥ 0.67 ; actual wRVUs.

†FPSC method in blue: data included only for cFTE ≥ 0.60 ; normalized wRVUs.

‡All faculty (in black) cFTE > 0 actual wRVUs.

For all the remaining analyses, we found that we had a few (1-4 pathologists) with reported days on service of 300 or greater. We excluded these pathologists from further analysis. The association between cFTE and days on service was more scattered for the hybrid practice model (Figure 6A). However, the association between wRVUs and days on service had a

picture similar to that of the general practice group: the densest cluster was between 100 and 150 days and wRVUs around 3000 and 6000 (Figure 6B). This suggests that the days on service may be useful in analyzing productivity for this group as well. However, there is more variability in the model, which we postulate is due to the effects of different subspecialties.

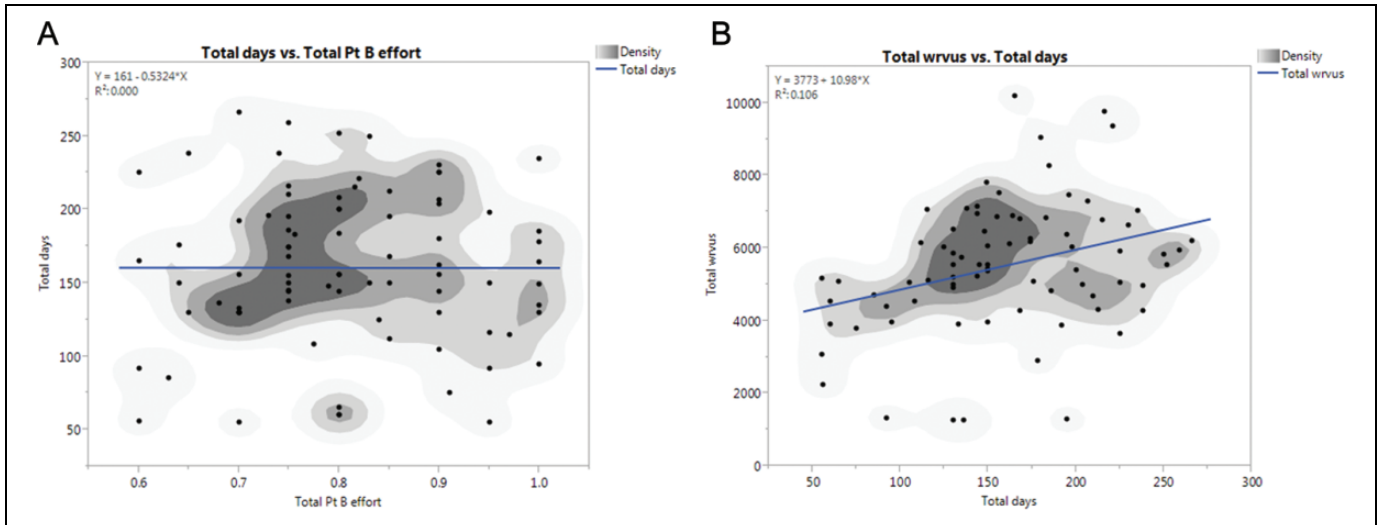


Figure 5. A, Scatterplot of days on service versus part B effort allocation for full-time (as defined by FPSC, cFTE ≥ 0.6) pathologists in the general practice model. B, Scatterplot of wRVUs versus days on service allocation for full-time (as defined by FPSC, cFTE ≥ 0.6) pathologists in the general practice model. The darkest color (top quartile) shows the densest cluster (top quartile) of points, each of which represents a faculty pathologist, while the lightest shows the least concentrated (bottom quartile). The intermediate colors show the second and third quartiles. The line of regression with its associated equation and coefficient of determination (R^2) are noted in the upper left corner. FPSC indicates Vizient-AAMC Faculty Practice Solutions Center; cFTE indicates clinical full-time equivalent; wRVUs, work relative value units.

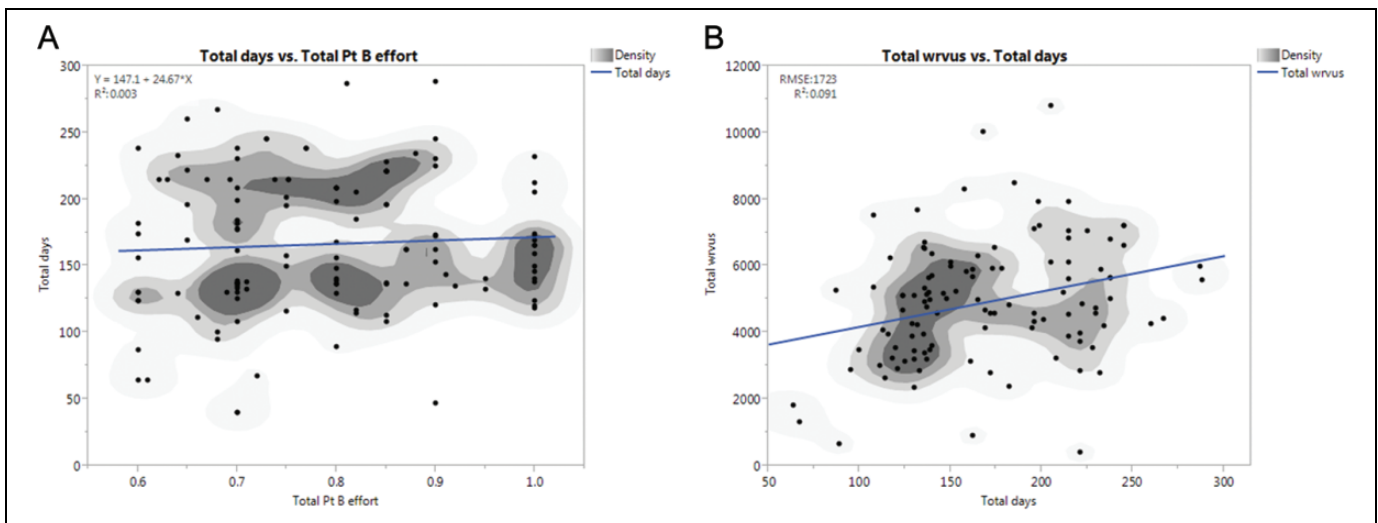


Figure 6. A, Scatterplot of days on service versus part B effort allocation for full-time (as defined by FPSC, cFTE ≥ 0.6) pathologists in the hybrid practice model. B, Scatterplot of wRVUs versus days on service allocation for full-time (as defined by FPSC, cFTE ≥ 0.6) pathologists in the hybrid practice model. The darkest color (top quartile) shows the densest cluster (top quartile) of points, each of which represents a faculty pathologist, while the lightest shows the least concentrated (bottom quartile). The intermediate colors show the second and third quartiles. The line of regression with its associated equation and coefficient of determination (R^2) are noted in the upper left corner. FPSC indicates Vizient-AAMC Faculty Practice Solutions Center; cFTE indicates clinical full-time equivalent; wRVUs, work relative value units.

For the single subspecialty practice model, a cluster analysis was once again difficult for both days on service and cFTE. However, the densest cluster was between 125 and 225 days and wRVUs between <1000 and 3500 (Figure 7A and B). This subspecialty practice model is quite different from the other models with lower wRVUs, at least for the subspecialties collected in our survey, and great variability. In Figure 8A, the box plot demonstrates wRVUs per single subspecialty for the 5

subspecialties with at least 10 FPSC-defined full-time (ie, ≥ 0.60 cFTE) pathologists. The median values differed significantly ($P < .0003$), with dermatopathology having the highest median wRVUs (7651; $n = 15$), followed by renal pathology (4018 wRVUs; $n = 10$), cytopathology (3395 wRVUs; $n = 15$), pediatric pathology (2486 wRVUs; $n = 13$), and neuropathology being the lowest (1820 wRVUs; $n = 15$). Although the difference among these subspecialties was statistically

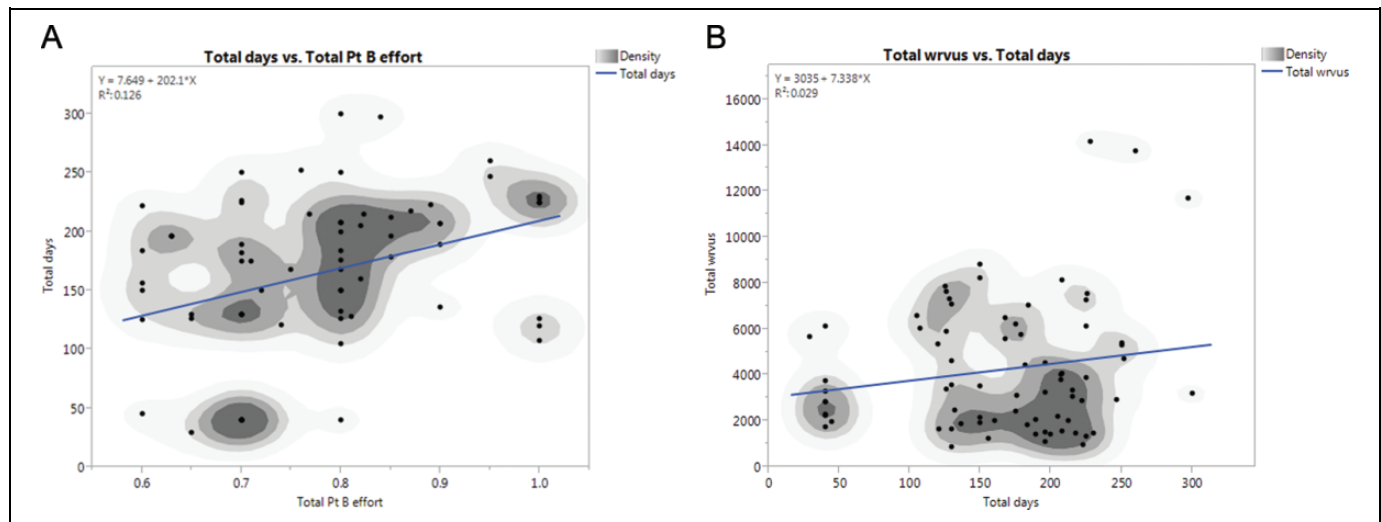


Figure 7. A, Scatterplot of days on service versus part B effort allocation for full-time (as defined by FPSC, cFTE ≥ 0.6) pathologists in the single subspecialty practice model. B, Scatterplot of wRVUs versus days on service allocation for full-time (as defined by FPSC, cFTE ≥ 0.6) pathologists in the single subspecialty practice model. The darkest color (top quartile) shows the densest cluster (top quartile) of points, each of which represents a faculty pathologist, while the lightest shows the least concentrated (bottom quartile). The intermediate colors show the second and third quartiles. The line of regression with its associated equation and coefficient of determination (R^2) are noted in the upper left corner. FPSC indicates Vizient-AAMC Faculty Practice Solutions Center; cFTE indicates clinical full-time equivalent; wRVUs, work relative value units.

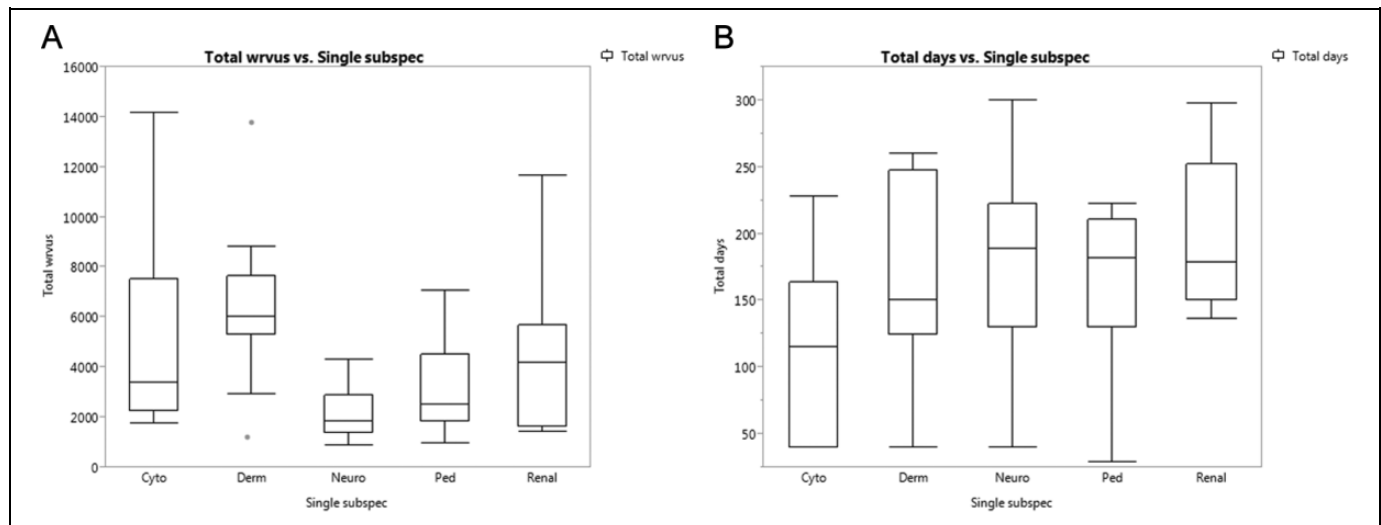


Figure 8. A, Box plot with outliers for wRVUs for 5 single subspecialties for full-time pathologists (FPSC definition of ≥ 0.6 cFTE). B, Box plot with outliers for days on service for 5 single subspecialties for full-time pathologists (FPSC definition of ≥ 0.6 cFTE). Only those subspecialties with at least 10 pathologists were included. FPSC indicates Vizient-AAMC Faculty Practice Solutions Center; cFTE indicates clinical full-time equivalent; wRVUs, work relative value units.

significant ($P < .0003$; Wilcoxon Kruskal-Wallis test), the number of pathologists in each group is small. The days on service also demonstrate variability in these subspecialties (Figure 8B). The remaining single subspecialties did not have enough pathologists (3 or fewer) for meaningful analysis. The multiple subspecialties practice model also demonstrated great variability with several clusters (Figure 9A and B). In this group, days on service would be difficult to be used as a marker for practice. We theorize this is again due to the multiplicity of subspecialty combinations with resultant different service obligations for days on service and resultant wRVU differences.

Based on the postulation that effort allocation for nonclinical effort (such as educational) might widely vary, and that effort allocation for nonclinical activities would have major effects, we reviewed how effort allocation differed among departments for residency program directors, since this is a defined role with similar responsibility across departments. In Figure 10, a scatterplot shows educational effort allocation (eFTE) as compared to the number of residents for residency program directors who were anatomic pathologists ($n = 29$ with 9 missing either an educational effort allocation or the number of residents). Although there was clustering and a

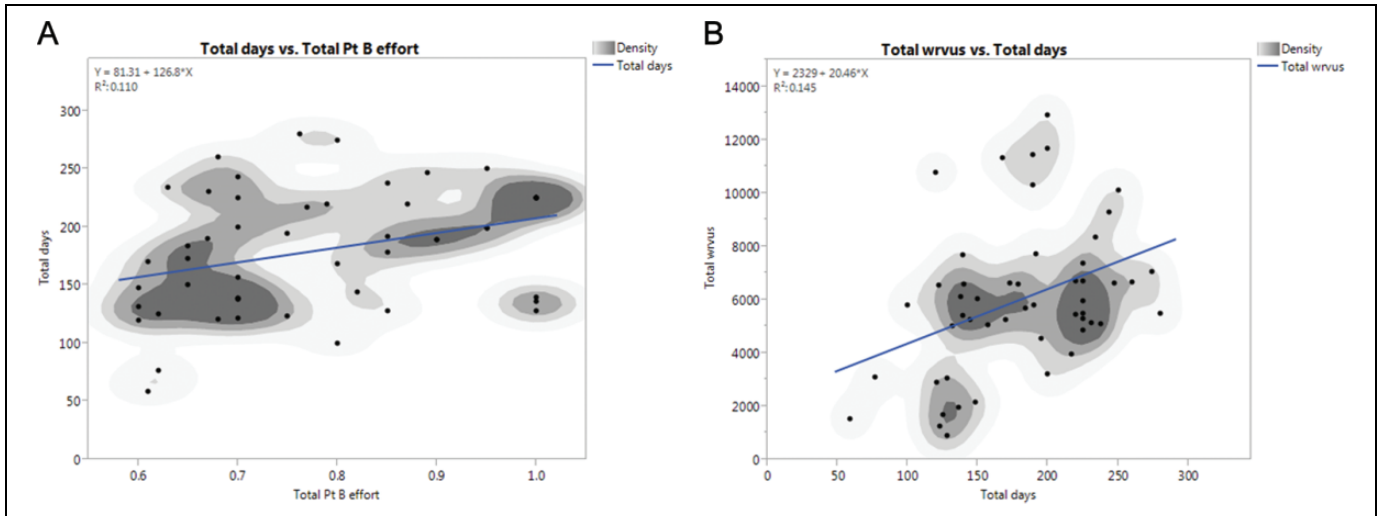


Figure 9. A, Scatterplot of days on service versus part B effort allocation for full-time (as defined by FPSC, cFTE ≥ 0.6) pathologists in the multiple subspecialties practice model. B, Scatterplot of wRVUs versus days on service allocation for full-time (as defined by FPSC, cFTE ≥ 0.6) pathologists in the multiple subspecialties practice model. The darkest color (top quartile) shows the densest cluster (top quartile) of points, each of which represents a faculty pathologist, while the lightest shows the least concentrated (bottom quartile). The intermediate colors show the second and third quartiles. The line of regression with its associated equation and coefficient of determination (R^2) are noted in the upper left corner. FPSC indicates Vizient-AAMC Faculty Practice Solutions Center; cFTE indicates clinical full-time equivalent; wRVUs, work relative value units.

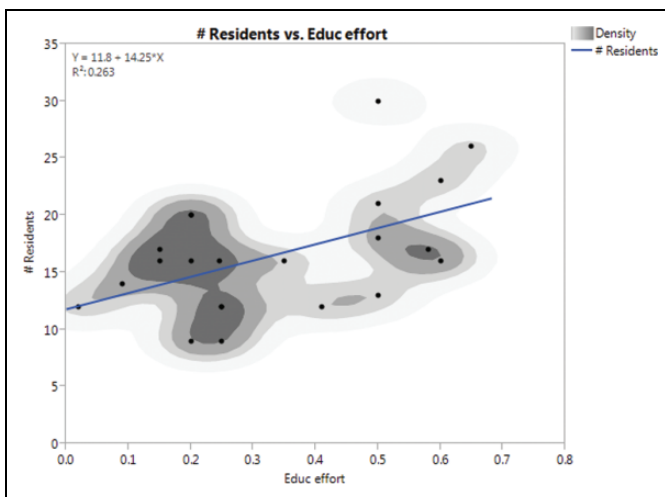


Figure 10. Scattergram of number of residents versus educational effort allocation for residency program directors within the anatomic pathologist data set. The darkest color (top quartile) shows the densest cluster (top quartile) of points, each of which represents a faculty pathologist, while the lightest shows the least concentrated (bottom quartile). The intermediate colors show the second and third quartiles. The line of regression with its associated equation and coefficient of determination (R^2) are noted in the upper left corner.

trend, there was also considerable heterogeneity—as an example, for a residency of 12 residents, 1 program director got 0.02 eFTE while another got 0.5 eFTE. Since the clinical effort allocation is the total FTE minus all other effort, theoretically the first program director could be 0.98 cFTE, but the second could theoretically be 0.5 cFTE while performing identical tasks.

Discussion

Study Value and Limitations

When we analyzed our data using MGMA and FPSC methodology, our median values were reasonably concordant (with a difference $<10\%$). This validates our data set as a comparison to other benchmarks. It also illustrates the need for transparency in understanding the methodology before applying it to departments or individual faculty members. Using actual data in a scatterplot analysis also gives the most information and can be used to highlight individual departments for comparison to their peers. As we accumulate years of similar data, we will be able to analyze the reasons for historical trends. Furthermore, we can use our data set to enter into discussions with these groups with the aim of improving their data.

Our study has the limitations of any survey. We cannot determine whether our data for 51 programs is representative of all academic departments. Clinical effort allocations are self-reported for all institutions in all the surveys. MGMA, FPSC and APC use self-reported data on clinical effort allocation and thus all methodologies are subject to the differences in self-reporting between programs.

MGMA Survey Value and Limitations

The MGMA survey includes both academic and community pathologists; and if faculty pathologists are mostly in clinical service roles, this may be a more appropriate comparison. But this benchmark excludes nearly half of academic pathologists. By lumping together full-time faculty without further correction, the MGMA benchmark avoids skewing the data by clinical effort. However, this may undervalue

productive faculty at the lower end of effort allocation while overvaluing unproductive faculty at the higher end. As an example, 2 faculty members at the median benchmark, 1 with a “true” effort of 0.67 cFTE and the other with a true effort of 1.0 would be considered equally productive, even though in this model, the lower effort pathologist is truly more productive.

FPSC Data Value and Limitations

The FPSC median benchmark is higher than other benchmarks because of wRVU normalization. Additionally, FPSC uses a gap-filling methodology and also uses an extended set of wRVU assignments by purchasing a supplemental source called the Complete RBRVS by Relative Value Studies, Inc. These differences also contribute to higher wRVU totals. The advantages of the FPSC approach are that it captures a slightly greater percentage of the overall data set, with the exact percentage depending on the practice model and it does allow for differences among faculty members with different effort allocations. In contrast to the MGMA data set, with the FPSC methodology the converse might be true: pathologists of equal wRVU productivity but with a lower assigned cFTE will appear more productive than pathologists with a higher assigned cFTE, even if they are actually on service the same amount of time. This reinforces the need for those involved in applying benchmarks at the departmental level to understand the process. FPSC collects its data from the billing records of the institutions with little departmental input, not from departmental self-reporting. However, the cFTE is collected from departments by a survey.

Value and Limitations of wRVUs as a Benchmark

Although wRVUs are an imperfect productivity measure in pathology, they are the accepted standard and are better than simply using case accession numbers, given the differences between the effort and time for various CPT codes. The wRVU is standardized for each CPT code, and the limitations of the wRVU system are the same across pathology departments, so that it allows effective comparisons for overall work volume. Cheung et al have proposed alternative workload measures that might ideally be more representative of the time and effort involved in pathology services.⁴ Given that reimbursement is determined by CPT codes which are tied to wRVUs and that wRVUs are an external measure that is used across all specialties, wRVUs were the productivity metric used for this study and are likely to be used by the administrators of academic health-care centers in the foreseeable future despite the problems we have demonstrated.

Although wRVU benchmark data may be somewhat useful for the entire pathology departments, we believe that our data demonstrate the inherent flaws in using wRVUs as a granular benchmark for individual faculty members. Any benchmark, whether MGMA or FPSC derived, comes with the same significant limitations as shown in our study.

First, wRVUs differ significantly across practice models. This is most striking for the single subspecialty model with the lowest median wRVU values. The low wRVU median in this group is likely due in part to the subspecialties collected in our survey; however, this results from a robust data set of 629 anatomic pathologists and is most likely representative of practice patterns throughout many academic departments. Even AP and SP are not well defined. It is probable that some pathologists in the general practice group practice only SP, whereas others practice SP, cytopathology, and autopsies. These differences, especially autopsies that have no wRVU, will tend to create differences. Finally, there are differences between median wRVUs for various specialties, but the numbers of pathologists available for analysis decrease as smaller and smaller subsets are analyzed.

The second major challenge is clinical effort allocation that varies significantly among departments and practice models. Without a better definition for cFTE, this is unlikely to change. At least for the general and hybrid practice models, days on service might be a useful marker, but this appears less useful for the single subspecialty and multiple subspecialties models. These differences are likely due to the differences in types of service assignments across the broad range of AP specialties as well as the intensity of services, both of which differ substantially from institution to institution. So some pathologists might have a lighter service day but work more days while others work more intensely for fewer days. Possibly some combination of hours per day and days on service (looking more closely at hours per week) will help to further define clinical effort. The APC is currently working with FPSC, and we will compare our analyses to determine if such an approach is feasible. Since effort allocation for education and administrative effort also varies significantly, another improvement might result from more robust guidelines to make such allocations more uniform and reproducible.

A final disadvantage of wRVUs in pathology include that many clinical activities in anatomic and clinical pathology do not have wRVUs, as these are not part B services that are directly reimbursed by the Centers for Medicare and Medicaid Services (ie, all part A activities).⁵ Finally, there can still be considerable differences in the time and effort required to sign out different cases with the same CPT code.

Summary and Conclusions

The use of wRVUs as a benchmark has many inherent limitations; nonetheless, it has been adopted by academic medical centers as a measure of productivity. The granular use of wRVU benchmarks applied to individual faculty members for targets, incentives, and compensations is very problematic, given the inherent limitations of wRVU data. Although applying the benchmarks to departments to entire departments for staffing analysis may be somewhat more useful, careful consideration will need to be given to practice models. Effort allocation influences productivity irrespective of the method of analysis used; better methods for determining effort and/or

time on service would help to improve productivity benchmarking. Nevertheless, this study demonstrates that the greater source of variability in data reporting is in effort assignment and reported days on service. Improving standards for such assignments and reporting will be of value in future efforts to track the clinical productivity of pathologists in academic departments of pathology.

Sources and Manufacturers

- a. Version 14.0, Microsoft Office Professional Plus 2010, Microsoft Corporation, Redmond, Washington.
- b. Version 11.0, SAS, Cary, North Carolina.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Note

I. Vizient, Inc., the largest member-driven health care performance improvement company in the country, provides innovative data-driven solutions, expertise and collaborative opportunities that lead to improved patient outcomes, and lower costs. The Association of

American Medical Colleges, a not-for-profit association representing all 145 accredited U.S. and 17 accredited Canadian medical schools, nearly 400 major teaching hospitals and health systems, and more than 80 academic societies, is dedicated to transforming health care through innovative medical education, cutting-edge patient care, and groundbreaking medical research.

References

1. Hsiao WC, Braun P, Becker ER, Thomas SR. The Resource-Based Relative Value Scale. Toward the development of an alternative physician payment system. *JAMA*. 1987; 258: 799-802.
2. Hsiao WC, Braun P, Yntema D, Becker ER. Estimating physicians' work for a resource-based relative-value scale. *N Engl J Med*. 1988; 319: 835-841. doi: 10.1056/NEJM198809293191305.
3. Thorwarth WT Jr. From concept to CPT code to compensation: how the payment system works. *J Am Coll Radiol*. 2004; 1: 48-53. doi: S1546-1440(03)00020-6.
4. Cheung CC, Torlakovic EE, Chow H, Snover DC, Asa SL. Modeling complexity in pathologist workload measurement: The automatable activity-based approach to complexity unit scoring (AABACUS). *Mod Pathol*. 2015; 28: 324-339. doi: 10.1038/mod-pathol.2014.123.
5. Robboy SJ, Weintraub S, Horvath AE, et al. Pathologist workforce in the united states: I. development of a predictive model to examine factors influencing supply. *Arch Pathol Lab Med*. 2013; 137: 1723-1732. doi: 10.5858/arpa.2013-0200-OA.