POMS PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT POMS

Vol. 27, No. 10, October 2018, pp. 1795–1815
ISSN 1059-1478 | EISSN 1937-5956 | 18 | 2710 | 1795
DOI 10.1111/poms.12835

Leveraging Big Data to Balance New Key Performance Indicators in Emergency Physician Management Networks

Krista Foster*
Katz Graduate School of Business, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA, Kmf88@pitt.edu

Pooja Penninti
Dietrich School, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA, ppennint@andrew.cmu.edu

Jennifer Shang
Katz Graduate School of Business, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA, shang@katz.pitt.edu

Sunder Kekre
Tepper School of Business, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA, skekre@cmu.edu

Gajanan G. Hegde
Katz Graduate School of Business, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA, hegde@pitt.edu

Arvind Venkat
Department of Emergency Medicine, Allegheny Health Network, Pittsburgh, Pennsylvania 15212, USA,
US Acute Care Solutions, Canton, Ohio 44718, USA, avenkat@ahn-emp.com

Managing emergency physicians is a complex task and has increasingly intensified with the recent consolidation of many emergency departments (EDs). Large-scale physician groups are facing challenges in resource deployment and performance evaluation. To objectively evaluate physicians across facilities, we leverage big data from an emergency physician management network and propose data-driven metrics using a large-scale database consisting of 84 hospitals, 1,079 physicians, and 10,615,879 patient visits in 14 states over 600,000 clinical shifts from 2010 to 2014. To ensure physicians are fairly evaluated and compensated within diverse facilities, we propose an index system and use clustering to help identify factors which might impact physician performance. The proposed indices benchmark physicians from the perspectives of *revenue potential*, *patient volume*, *patient complexity*, and *patient experience* by controlling for exogenous factors at the facility level. We empirically show the *volume* and *complexity* indices are key elements of the *revenue potential* index, and use two-stage least squares regression to relate *volume* and *complexity* and uncover their drivers. *Revenue potential* and *patient experience* are found to be positively correlated, which suggests productive physicians are often liked by their patients. Through implementing the proposed evaluation system, administrators can better manage and incentivize physicians and provide directions for performance improvement, while controlling for location idiosyncrasies. The proposed framework can also be adapted to non-medical professional settings such as value chains, where employees often provide services in various profit- and cost-centers.

Key words: big data analytics; physician performance; patient satisfaction; managing value chain networks
History: Received: June 2017; Accepted: December 2017 by Kalyan Singhal, after 1 revision.

1. Introduction

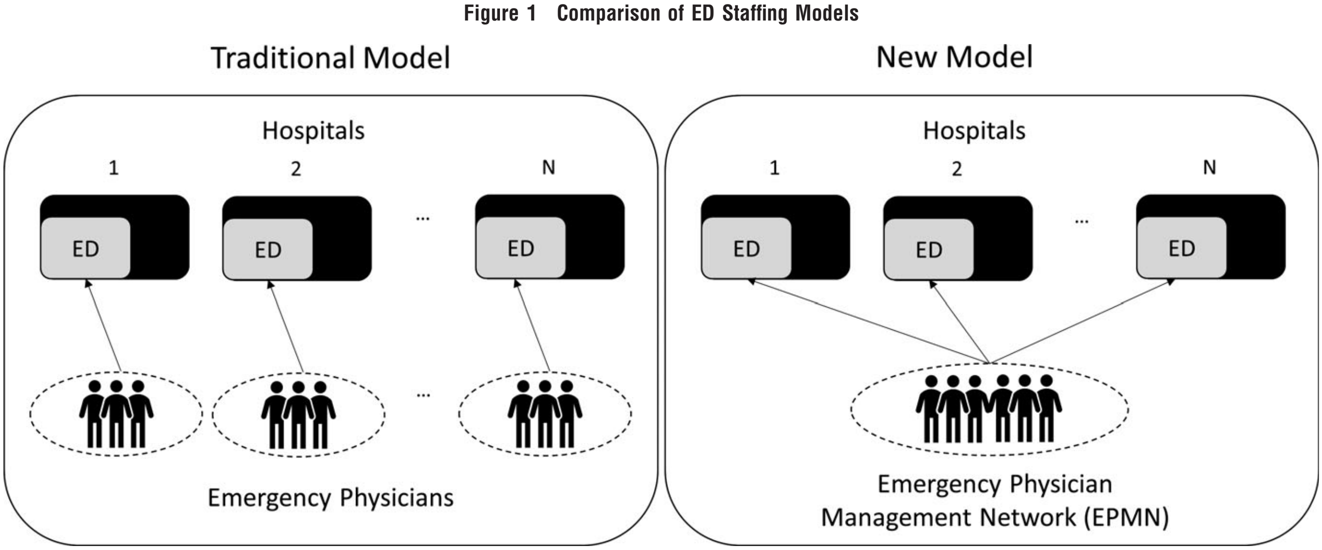
"The director of an emergency medicine group is struggling. The group of emergency physicians has provided emergency department (ED) services to a five-hospital health system for the past ten years. Given new legislation, the health system is seeking opportunities to maximize ED productivity and minimize costs. Simultaneously, the emergency physicians in the group are in revolt. They feel that based on the particular hospitals in which they work; their productivity-based compensation varies drastically – driven by factors out of their control such as location, services available at each hospital, and the scheduling idiosyncrasies of each hospital. The emergency physicians also feel that the health system is imposing unrealistic productivity and patient experience goals upon the group without any objective basis for measuring their performance against peers. The director must address the demands from the health system to show where productivity can be improved and from the physicians to measure their productivity objectively. Now the director is at risk of seeing physicians leave, losing the contract with the health system or both. Is there a better way?"

The emergence of big data, typically defined by its volume, velocity, and variety, is transforming the health care industry. Massive investments in electronic health records and advances in technology such as smartphones and wearable medical devices have given rise to both structured and unstructured health care data, which are continuously changing. Such innovations have created numerous opportunities for both researchers and practitioners, and rich data is being collected and utilized by health care providers, pharmaceutical companies, and insurance companies. For instance, several organizations have launched big data initiatives in health care, ranging from efforts to improve medical treatments to personalized medicine (Nambiar et al. 2013). In many cases, the goal is patient-centric health care (Sonnati 2015). Rising health care costs have motivated other projects as well. Improvements in record-keeping have helped to reduce health care costs due to fraud, abuse, waste, and erroneous insurance claims (Srinivasan and Arunasalam 2013). As health care systems aim to balance the patient experience (or patient satisfaction) with costs and revenues, data from electronic health records and RFID tags have become invaluable. The increasing availability of data relating to both patients and physicians provides an opportunity to reevaluate the methods used to measure and assess physician performance.

This research was motivated by the common dilemmas facing executives in managing consolidated multi-facility emergency physician management networks (EPMNs). This group of emergency physicians (EPs) has provided emergency department (ED) services to multi-hospital health systems for several years. The increasing consolidation in staffing emergency departments is both an adaptation to the changing health care landscape and a mechanism to remain competitive in an era of shrinking profit margins in providing patient care. Previously, single hospitals would either employ directly or contract with a small group of emergency physicians to provide patient care services in their ED. Natural growth might lead this hospital to partner with one or more local hospitals and therefore hire more emergency physicians or ask their contracting local group to do the same.

With rapid consolidation of hospitals into health care networks to achieve economies of scale, the nature of staffing EDs has changed (see Figure 1). These changes have resulted in horizontal integration in terms of staffing as well as risk-pooling. Many health care systems today have multiple hospitals dispersed geographically (similar to a distributed supply chain), have varying profit margins, and are of different appeal to an already undersized (short supply) emergency physician workforce (Reiter et al. 2016). To address these challenges, health care systems often seek large EPMNs to manage their EDs and assume

Figure 1 Comparison of ED Staffing Models



the financial responsibility of their emergency physicians' compensation.

Multi-hospital ED consolidation through EPMNs helps aggregate resources and provides adequate capacity to meet this need. However, as the demand for physician services is growing faster than supply, EPMNs are constantly under pressure to maintain physician supply to meet patient demand. Despite the varying level of profitability under different hospitals, the administrators need to hire quality physicians to work, especially in less appealing ED facilities and locations. They also need to ensure physicians are fairly assessed and compensated. Herein lies a typical value chain challenge of maintaining a sufficient level of resources (quality physicians) and efficiently managing its operations (ED) where care must be available at all times to meet patient demand under budget constraints. There is an earnest need to explore this evolving industry and identify effective ways to deliver care effectively and efficiently.

We have partnered with a physician-owned EPMN that serves about six million patients per year at over 170 sites in 21 states as of 2017 and continues to grow. The EPMN under study has surveyed and documented emergency medicine (EM) trends and observed noticeable increases in the number of EDs managed by physician networks. Specifically, nearly 40% of all EDs in the United States are currently managed by EPMNs. These complex networks constantly strive to balance the demands of three key stakeholders: the company, the facilities and their patients, and the physicians. While the company aims to acquire new contracts with facilities and attract physicians, every facility must be staffed, and physicians desire to be paid competitively. As EPMNs normally contract with multi-facility health care systems, management constantly faces challenges when staffing EDs in hospitals of various performance levels (e.g., productivity and patient experience), while under the pressure to ensure all physicians are evaluated and compensated fairly. With continuing health care reform efforts and the current shortage of EPs, the EPMN under study is constantly seeking methods to incentivize and retain current physicians, attract new physicians, and engage in continuous improvement to enhance ED performance.

To maintain contracts with health care systems, efficiently manage EDs, and ensure physicians' continued dedication to their practice, EPMNs must focus on both clinical and operational performance. Clinical performance (e.g., clinical outcomes, unexpected return visits, medico-legal risk) depends heavily on the quality of the physicians hired and available risk management and clinical education programs. Therefore, clinical outcomes are largely subsumed by hiring practices, which is beyond the scope of this research. However, operational metrics are important to maintaining health system contracts and physician investment in the EPMN, and this is an area that is largely unexplored in the emergency medicine and related operations literature. How to fairly and effectively evaluate and incentivize EPs while improving productivity and patient experience has become a pressing issue for the executives. Thus, management is in search of a method that would equitably assess physicians within the EPMN based on their productivity (objective score) and patient experience (subjective score) in order to balance performance and customer satisfaction, both of which are key to maintaining health system contracts for ED services. Their relative standings among physicians in the network will help management to make justifiable decisions on performance-based compensation and training requirements.

As variation in productivity has implications on ED patient flow and waiting times, objective measurement of EP performance is important, particularly when it comes to its consequent impact on ED performance. Researchers studying EP productivity, however, have mostly focused on a single facility. In this research, we address the emerging trend that emergency physicians work in multiple facilities. We conduct a multi-center, multi-year, and multi-physician study to investigate EP operational performance in the EPMN setting to draw lessons on how to measure and enhance physician productivity and patient experience. For such a study, we make use of EP profiles, daily schedules, monthly patient experience surveys, patients' visit details, and patients' insurance information to better understand the performance of physicians in this network using big data analytics.

EPs encounter a number of work conditions that affect their performances. These circumstances make the revenue generated by an EP less relevant to performance evaluation. For instance, if patients are uninsured, they cannot pay for treatment, and different insurances pay different amounts for similar services, completely out of the treating physician's control. Moreover, an EP's output flow rate (Patients/hr) is highly dependent on patient acuity and chronic medical conditions, the patient arrival rate, and the ED's capacity. As an example, an elderly adult patient with chest pain and several comorbidities will likely take longer to treat than a young adult with an ankle injury who is otherwise healthy, as the former is much more serious and complex. It would be unfair to penalize physicians who treat complex patients for their resultant lower patient flow rates, or equivalently, to reward physicians who are assigned to simpler cases by relying solely upon the volume-based Patients/hr measure. Relative value units (RVUs) are thus often used to assess emergency physician performance as they mirror the time and supplies/devices needed from the health care workers/facility to care for the patient as well as the cognitive expertise and potential medico-legal risk (Venkat et al. 2015; see Technical Supplement S.1). RVUs reflect the revenue potential for a particular visit, while the actual revenue generated depends on the specific insurance, location and type of hospital. Conventionally, RVUs/hr is a marker for revenue potential and is a proxy for physician productivity, capturing the yield from both volume and complexity of a physician's efforts. However, one should note that RVUs/hr is not a pure measure of productivity nor is it an absolute measure of revenue.

To date, comparison of EP performance within such a large network has not been available due to the technical difficulties involved (e.g., demand heterogeneity in different facilities, shift variation, and varying availability of diagnostic equipment). RVUs/hr is driven by complexity (RVUs/Patient) and patient flow (Patients/hr) at the facility level, both of which affect an EP's performance. Thus, significant disparity may be found in terms of facility and physician performance. The variation across facilities thereby limits the efficacy of the conventional RVUs/hr measure, making it necessary to take into account these differences while developing fair performance metrics to assess physicians within the EPMN.

In order to overcome the aforementioned difficulties when evaluating EP productivity, we propose four new indices for assessing physician performance relative to peers: *revenue potential index*, *patient volume index*, *patient complexity index*, and *patient experience index*. We employ the revenue potential index and patient experience index within a large EPMN to differentiate the high performers from those physicians lagging behind on each metric, resulting in a 2-by-2 graph. Then, we segment physicians into clusters to uncover possible physician characteristics affecting their performance. We empirically verify that physician productivity is a function of the proposed complexity and volume indices and subsequently identify drivers of the two proposed indices such that management can help physicians target improvements in both dimensions. Our study highlights that the use of big data analytics to manage complex and large-scale physician groups has major potential for developing and deploying new metrics that are sophisticated, yet relatively straightforward to implement, and offers operational insights for volume and complexity. The proposed metrics are transparent and take into account facility-specific differences. These metrics can be linked to performance-based pay, making them attractive to physicians and therefore mitigating the obstacles management faces in recruiting and assigning physicians to facilities.

The remainder of this study is organized as follows. In section 2, we review the related literature. The theoretical framework, rationale for indices, and hypotheses are discussed in section 3. In section 4, we describe the data and variables and develop the indices. We segment physicians into clusters to derive insights in section 5. In section 6, we present the empirical results. Managerial implications, limitations, and conclusions are given in section 7.

2. Related Literature

Service outputs can be classified into two components: quantity-oriented (volume) and quality-oriented (process and outcome) (Grönroos and Ojasalo 2004). In EDs, patients (heterogeneous customers) arrive unannounced and expect high quality care (output) within a short time. The idea of service systems with heterogeneous customers is not new (Armony and Ward 2010, Ward and Armony 2013); and most of these systems focus on throughput (volume). Similarly in EDs, operational efficiency and patient experience are important to all stakeholders (e.g., patients, physicians, and administrators). Facilities and physicians are judged on patient experience via survey data, such as those from the Press Ganey© (PG) survey (see Technical Supplement S.1). While patient experience has become a focal point of health care providers, physician productivity remains crucial from a resource management perspective. Management's ultimate goal is to concurrently achieve high productivity and excellent patient experience. In this research, we aim to address these two aspects of performance while making use of the recent explosion of big data in the health care sector.

The impact of big data and analytics within the health care industry has been widely recognized by health care professionals and beyond. Murdoch and Detsky (2013) outlined ways in which big data may be used to improve the quality and efficiency of health care delivery, from better knowledge dissemination to personalized patient care. Raghupathi and Raghupathi (2014) cited specific examples in their review of big data in health care, including earlier predictions of sudden increases in flu-related emergency room visits. McKinsey & Company also weighed in on big data's place in the health care sector (Kayyali et al. 2013). Both Ellaway et al. (2014) and Moskowitz et al. (2015) discuss the impact of big data on education and training for clinicians and other health care professionals. Furthermore, Obermeyer and Emanuel (2016) recommend the development of algorithms to make the best use of health care data. Most recently, Ba and Nault (2017) identified health care IT that incorporates OM methods as an emerging research area. All of these researchers agree that the availability of big data, including physician work histories and patient records, is changing the management of facilities, physicians, and patients. EDs and physician management networks are no exception. For example, the American College of Emergency Physicians (ACEP) has introduced the Clinical Emergency Data Registry (CEDR), which is designed to collect data from emergency physicians across the United States to measure health care quality and to identify practice patterns, trends and outcomes in emergency care (ACEP 2016). The ultimate goal of CEDR is to inform emergency physicians and eventually improve the overall quality of emergency care, suggesting that emergency medicine clinicians are embracing data-driven knowledge and decision making.

Several operations management researchers have also focused on problems in the health care sector, leveraging data to motivate research questions and validate models and theories. The single-facility health care operations literature abounds, and EDs have been the center of several studies (KC 2013, Lee et al. 2015, Powell et al. 2012, Song et al. 2015). For instance, Venkat et al. (2015) used ED data to gain insights into its management and revenue generation. However, operations researchers who have studied multiple facilities mainly focus on differences in hospitals and compare their efficiencies (Bhargava and Mishra 2014, Blank and Eggink 2014, Büchner et al. 2016, Theokary and Ren 2011). Angst et al. (2011) studied how US hospitals convert existing medical technologies into integrated information technology. No research hitherto has examined the dynamics between physicians and patients, and compared individual practitioners' performances across the multiple EDs each physician serves. Also, no researchers to our knowledge have explored the implications of multi-facility physician management networks.

Different from the emergency medicine literature's physician studies (e.g., Brennan et al. 2007, Clinkscales et al. 2016, Johnson et al. 2008), we look into physician performance with the aim to efficiently manage and incentivize physicians within an EPMN that must staff many EDs and attract large numbers of practicing EPs. Our research differs significantly from both the operations and EM literatures as we consider two key dimensions of physician performance: patient experience and productivity (the latter comprising volume and complexity). Through big data, we are able to offer practice-based evidence to develop insights for multi-facility physician management networks instead of the conventional approach. The proposed metrics take into account physician efforts and peer effects. Our setting is unique as the physicians under study work at multiple sites under the management of an organization, whose goal is to ensure appropriate staffing across multiple facilities under contract for emergency services.

Our research benefits EPMNs, an increasingly prevalent EM practice model. Unlike the frequently used data from the Centers for Medicare and Medicaid Services (CMS), all of the physicians in our data are managed by an EPMN, and the proprietary data uniquely links patient, physician, hospital, revenue, and insurance information. To our knowledge, this is the first research to study physician productivity on this scale and with this level of detail. We contribute to the literature by addressing the trend in which physician work assignments include varying hours across multiple sites and shifts.

By developing relative physician performance indices, we help management understand the impact of multiple-facility employee assignments on both relative productivity and patient experience. The visual display of each physician's two-dimensional scores is conducive to employee development, as it allows management to readily assess physician performance.

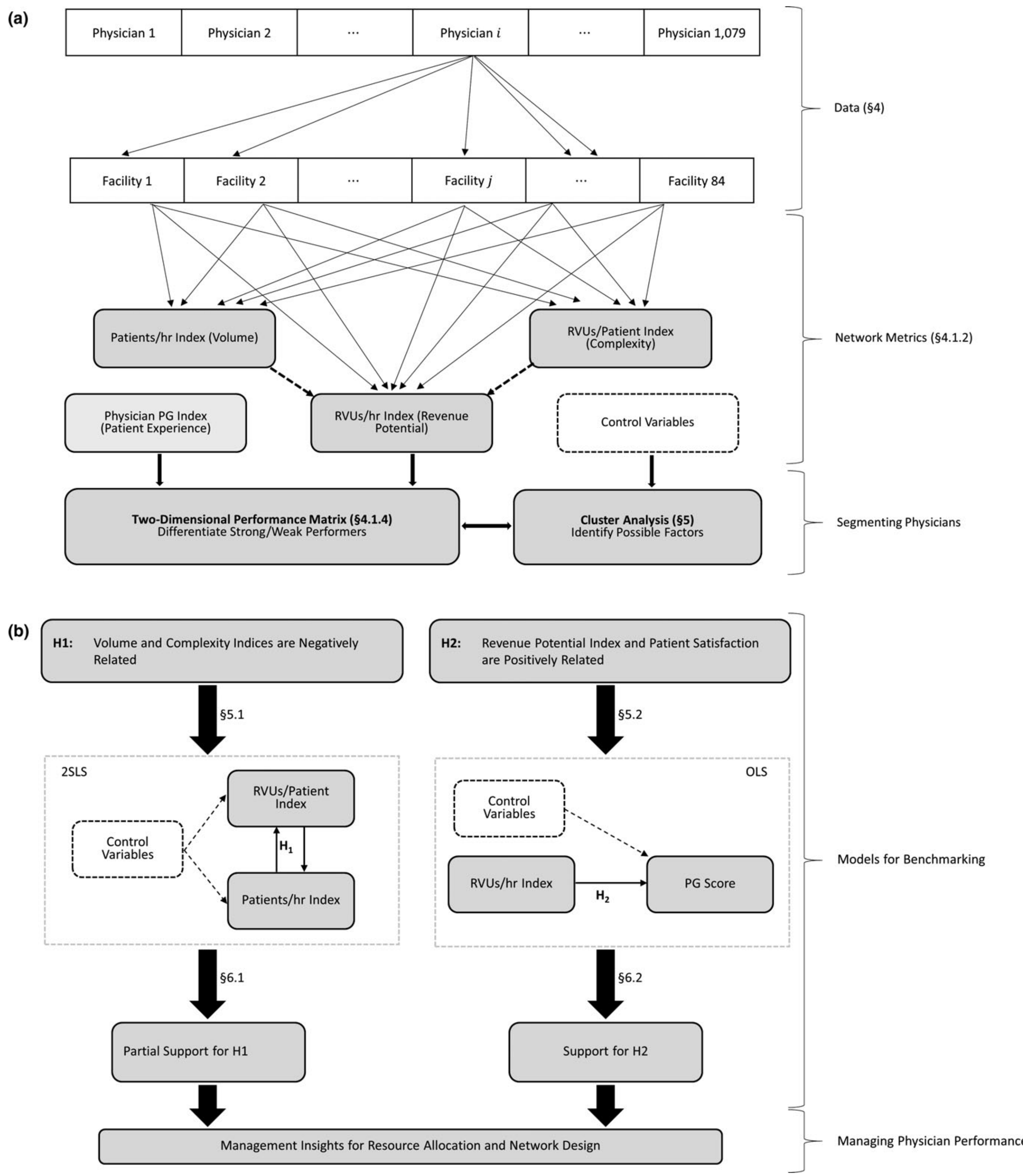
3. Theoretical Framework

Our research framework is motivated by management's need to explicitly and equitably differentiate strong and weak performers within the network, so as to provide justifiable bonuses or prescribe remedial actions. Through identifying driving forces contributing to physician performance differences, we offer the EPMN management implications for continuous improvement.

We match the 10,615,879 patient visits to physicians and facilities to develop indices that assess physicians within the network (Figure 2a). Although physicians may treat many patients across multiple facilities, the indices quantify performance within the network using a single value. Such metrics allow management to easily identify the best performers to reward with performance-based bonuses, whereas the poor performers would require intervention in the form of training. To gain managerial insights into effective network resource management, we conduct cluster analysis to sort physicians into distinctive groups.

Justification for the proposed indices, including a stylized example to illustrate their benefits, is detailed in section 4. We then use the indices to develop statistical models and benchmark physician performance (Figure 2b). Our research relates Physician's Revenue Generation Potential with Patient Volume and Patient Complexity, while controlling for patients' medical and demographic information as well as physician and facility characteristics.

Figure 2 (a) Developing Network Metrics for Physician Performance (b) Research Outline



Using the performance indices, we develop a two-stage least squares model to simultaneously reveal the underlying drivers of the volume and complexity indices. Subsequently, we link Physician's Revenue Generation Potential to Patient Experience to better understand how different aspects of performance interact and to empirically test if tradeoffs exist for physicians at the network level. By using these metrics and statistical models, we gain managerial insights for an EPMN.

3.1. Factors Impacting Physician Performance

We develop integrated relative performance indices to assess the EPMN's physicians in section 4, addressing management's need for a fair and coherent network performance metric. In this section, we develop models that relate these indices to physician, patient, and facility traits to better understand the underlying drivers.

3.1.1. Identifying the Drivers of Performance Indices and Linking the Indices. In order to identify factors driving physician performance, we employ two-stage least squares (2SLS) regression to simultaneously model the proposed relative performance metrics, RVUs/Patient Index and Patients/hr Index, by Equations (1)–(2). The model is estimated using log-transformed values of RVUs/Patient Index and Patients/hr Index to address the nonlinear relationship characterized by diminishing returns.

$$ln(RVUs/Patient\ Index) = \alpha_{11} + \beta_{11}\ ln(Patients/hr\ Index) + \gamma_1^T Controls + \varepsilon_1 \quad (1)$$

$$ln(Patients/hr\ Index) = \alpha_{21} + \beta_{21}\ ln(RVUs/Patients\ Index) + \gamma_2^T Controls + \varepsilon_2 \quad (2)$$

While patient complexity and patient flow are related (RVUs/Patient × Patients/hr = RVUs/hr) on a system-wide level, these have not been studied at the physician-level. For example, Eitel et al. (2010) found that at the ED-level, the criticality of patients is associated with flow. This relationship may also hold at the physician-level. Based on management's desire to measure physician operational performance on two dimensions (patient satisfaction and productivity), it is important to fully understand both metrics and their drivers (patient complexity and patient flow). As patients with high complexity require more physician time to treat, we posit that an inverse relationship exists between RVUs/Patient Index and Patients/hr Index (see the left-side of Figure 2b). Hypothesis 1 tests such a relationship.

H1. *There exists a negative relationship between the number of patients that a physician treats per hour relative to peers (Patients/hr Index) and the relative number of RVUs the physician generates per patient (RVUs/Patient Index). That is,* $\beta_{11} < 0$ *and* $\beta_{21} < 0$.

Although H1 seems intuitive, the unpredictable circumstances of emergency care may counterintuitively disrupt the presumed inverse relationship between complexity and volume. In emergency medicine, there are numerous examples where high RVU patients are treated quite efficiently in the ED. The ensuing examples justify the importance of testing H1. Consider a patient with a severe allergic reaction. The patient will be rapidly treated, observed, and commonly discharged, but the visit will likely result in high RVUs due to the critical consequences if this condition is not treated appropriately. In other cases, an EP may be constrained by the environment, such as the specific number and type of arriving patients or the availability of beds for admitted patients. For instance, there may be a high number of patients requiring admission to the hospital during flu season, but most of these patients do not require the resources of an intensive care unit (ICU). In this case, high RVU, critically ill patients may rapidly be placed in ICUs, while less complex patients requiring admission may wait hours for an inpatient bed. Such circumstances make the exploration of H1 highly relevant to the study of productivity in the emergency department.

3.1.2. Linking RVUs/hr Index to Patient Experience Index in the Network. After identifying the drivers of physician performance in Equations (1)–(2), we now explore the extent to which physician performance is related to patient experience as measured by the Press Ganey© (PG) patient experience index. We present an OLS model to describe the relationship between RVUs/hr Index and PG Index while controlling for patient and physician level differences in Equation (3).

$$PG\ Index = \alpha_{31} + \beta_{31}(RVUs/Hour\ Index) + \gamma_3^T Controls + \varepsilon_3 \quad (3)$$

As RVUs/hr consists of two components, Patients/hr and RVUs/Patient, which are presumed to be the primary drivers of RVUs/hr Index, we study how each of these components influence patient experience. First, we consider Patients/hr, which reflects the logistics of providing patient services. If an ED visit is extended due to a long wait time or prolonged service, the patient may blame the physician, leading to dissatisfaction. This frustration would be reflected in physician PG survey responses. Such phenomena have been studied in the EM literature. For example, Handel et al. (2014) observed that patients with low door-to-room times gave higher experience scores, and Pines et al. (2008) found a negative relationship between overcrowding (longer wait times, prolonged treatment times) and patient experience assessment. Furthermore, Hwang et al. (2015) found that implementation of a fast track significantly increased patient experience. Thus, we expect Patients/hr to be positively associated with patient experience.

Alternatively, Boudreaux and O'Hea (2004) found that the strongest predictor of ED patient satisfaction is how satisfied the patient is with interpersonal interactions with the ED staff. We conjecture that patients may view visits with higher RVUs (more tests and resources) as a sign that physicians care and are sincere (positive interactions). Patients may perceive increased numbers of examinations, procedures, or therapeutic interventions, leading to higher RVUs, as signs that they have been taken seriously, and thus RVUs/Patient may reflect a physician's individualized focus toward patients. Thus, RVUs/Patient will also be positively correlated with patient satisfaction, even though RVUs/Patient and Patients/hr are inversely related.

Together, we anticipate that as a physician's productivity (RVUs/hr = RVUs/Patients × Patients/Hr) increases, patients will perceive that the physician is both highly skilled and focused, leading to higher PG percentile rank scores (hereafter abbreviated as PG scores), resulting in a higher PG Index. Boudreaux et al. (2006) further provided the theoretical foundation that patient satisfaction is dependent on physician performance. Therefore, as a physician's ability to speed up processing or handle more complex patients is associated with high RVUs/hr Index, we posit that physicians with high RVUs/hr Index will exhibit better PG scores relative to their peers (see the right-side of Figure 2b).

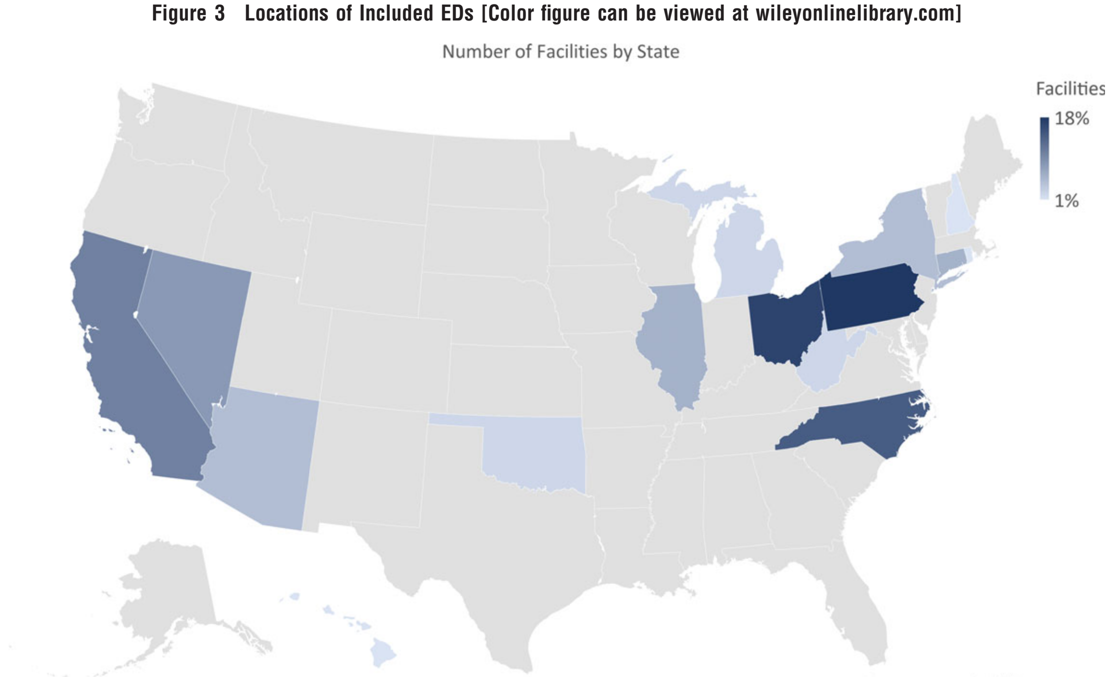
H2. *There exists a positive relationship between the number of RVUs that a physician generates per hour relative to peers (RVUs/hr Index) and the relative patient experience score (PG Index). That is,* $\beta_{31} > 0$.

4. Data and Variables

The data under study were collected from one of the largest EM providers in the United States. These data are unique as they are from a large private EPMN, which includes numerous physicians across 84 facilities in 14 US states from 2010 to 2014. The percentage of facilities in each state is shown in Figure 3. Our data initially contained 90 EM facilities, 1434 physicians, 622,716 clinical shifts, and 11,060,222 patient visits, comprising visit records, physician profiles, health care provider schedule logs, and facility profiles from January 2010 to June 2014. Visit characteristics, including Current Procedural Terminology Evaluation and Management (CPT E/M) codes and relative value units (RVUs) generated, were abstracted by trained coders. The coders need to acquire relevant certification(s) between their second and third employment year, with ongoing training, auditing, and external evaluation. The EPMN also maintains a demographic and credentialing database of all physicians. Physicians' clinical hours were tracked electronically using physician scheduling software (e.g., Tangier; Sparks, MD). Patient experience data (Press Ganey Associates Inc., South Bend, IN) were linked to physicians monthly.

These data are stored in various large databases, four of which are used in this research. The digitally

Figure 3 Locations of Included EDs [Color figure can be viewed at wileyonlinelibrary.com]



recorded databases include a variety of data types (numeric, factors, strings, dates), and as might be expected, not all data are initially stored in the correct format (e.g., numeric displayed as string). Similarly, date-time information for visits and shifts do not follow a standard format and required careful conversion. In addition, several measures (e.g., patient length of stay and physician shift length) need to be computed from the database, and missing data needed to be removed. During each stage of the analysis, we applied inclusion/exclusion criteria and verified the real-world validity of descriptive values by consulting with our clinical collaborator.

In the initial stages, we ensure that these four databases are correctly linked using common fields. We then use the matched data to construct shift-level and month-level data, which are combined to obtain a record for each physician in each facility. This procedure involves the calculation of several aggregate variables, such as the proportion of hours worked during each shift and the Advanced Practice Provider (APP) (e.g., physician assistant or nurse practitioner) support ratio. These two measures present challenges as physician and APP schedules constantly change, and each shift needed to be broken down into hourly segments to determine these ratios. The initial task was demanding, but we are able to automate the procedure for future use by developing user-defined functions. The final step was to use the derived data to compute the indices and independent variables for each physician. The resulting dataset was used for statistical analyses herein.

Six newly acquired facilities and 355 physicians with short work histories (<500 patient visits) in the dataset were excluded. Thus, the final data for this study include 84 EM facilities, 1079 physicians, and 10,615,879 patient visits with a total of 54 attributes. The facilities, which vary in yearly volume and capabilities (e.g., trauma designation or academic), are primarily located in urban and suburban regions, with only five EDs (6%) serving rural communities. Due to the range of locations and facility types, some EDs are profit-centers, while others are cost-centers, but the EPMN is still responsible for maintaining physician staffing levels at all facilities. Table 1 includes summary statistics for the final attributes we compiled for each physician and used for our analysis, including facility characteristics, patient visit records, the merged information from physicians' and patients' interactions, and physician's demographic and professional attributes. The data used for the analyses herein were aggregated over the entire study period, resulting in one record per physician (n = 1079). However, the same methodology could be applied using monthly data.

Table 1 Summary Statistics (10,615,879 patient visits of $n = 1079$ physicians at 84 facilities)

Level	Categorical variable	Count (%)
Facility characteristics	≤20,000 Yearly Visits	23 (27.4%)
	20,000–40,000 Yearly Visits	27 (32.1%)
	40,000–60,000 Yearly Visits	22 (26.2%)
	60,000–80,000 Yearly Visits	8 (9.5%)
	80,000–10,000 Yearly Visits	3 (3.6%)
	>100,000 Yearly Visits	1 (1.2%)
	Teaching Facility	15 (17.9%)
	EM Residency Training Site	12 (14.3%)

Level	Continuous variable	Mean	SD
Visits	Average Patient Age	40.406	9.060
	% Male Patients	44.274	3.130
	% ICD9 Group 1	36.620	5.515
	% ICD9 Group 2	36.539	8.801
	% ICD9 Group 3	23.404	5.300
	% Admitted	17.673	9.980
	Commercial Index	1.004	0.063
	Medicaid Index	0.992	0.184
	Medicare Index	1.018	0.312
	Self-Pay Index	0.991	0.109
Physician	Physician Age	45.870	9.994
	# Facilities Worked	2.456	2.294
	% 6AM-3PM Hours	36.434	14.337
	% 3PM-12AM Hours	47.467	11.609
	% 12AM-6AM Hours	16.086	14.998
	Coding Com/Patient	0.507	0.484
	Physician PG Scores	54.241	24.773
	Physician PG Index	0.986	0.443
	APP Support Ratio (%)	22.772	10.515
	RVUs/hr	9.623	2.198
	RVUs/Patient Index	1.005	0.089
	Patients/hr Index	0.995	0.235
	RVUs/hr Index	0.991	0.188

Level	Categorical variable	Percent
Physician Characteristics	Male	69.045
	White	78.221
	Primary Pediatric Practice	4.912
	Efficiency Training	29.935
	Patient Satisfaction Training	34.198

Notes: See Tables A1 and A2 in Appendix A for detailed variable definitions. Indexed measures are computed by Equations (7)–(10) and Equation (A1).

The patient visit records include the hospital, the date and time of patient arrival at and departure from the ED, age, gender, the attending physician, disease codes (ICD-9), and the discharge disposition (admitted, discharged, transferred, elopement, left without being seen, or died) for each patient visit. Additionally, information regarding the proportion of patients with each payment source (Commercial Insurance, Medicare, Medicaid, or Self-Pay) and the RVUs associated with the visit are provided. We grouped ICD-9 codes into three categories (Group 1: Circulatory, Respiratory, Digestive, and Genitourinary; Group 2: Symptoms, Signs & Ill-Defined Conditions; and Group 3: Injury & Poisoning) that reflect the most prominent diagnostic groups attributed to ED patients. The patient level variable definitions are detailed in Table A1.

Each attending physician has a profile, which denotes demographic information, date of residency completion, and if explicitly trained on efficiency and patient experience by the EPMN. In addition, the health care provider work logs contain data corresponding to physicians and APPs. For each shift worked, these include start and end times, the facility, whether a physician or APP, the provider's monthly PG patient experience percentile rank score, and the facility's monthly PG percentile rank score. Note that PG scores for physicians are reported on a monthly basis for each facility, so we use weighted averages to compute one physician PG score per physician at each facility, weighting by the number of hours worked during each month (Equation (4)). Within each facility j, the total number of months that physician i has worked is $S_{i,j}$ The total number of hours physician i worked during month l at facility j is denoted as $Hours_{i,j,l}$, and the corresponding physician PG score is $PG_{i,j,l}$. See Table A2 for definitions of all physician-level variables.

$$Physician\ PG\ Score_{i,j} = \frac{\sum_{l=1}^{S_{i,j}} PG_{i,j,l} \times Hours_{i,j,l}}{\sum_{l=1}^{S_{i,j}} Hours_{i,j,l}} \quad (4)$$

4.1. Physician Performance Measures and Indices Development

The EPMN seeks to explicably and impartially assess all physicians within the network. Management wants to reward the strongest performers, while improving the performance of the weakest performers. Physician performance is based on two dimensions: patient experience and physician productivity. The EPMN administers patient experience surveys (Press Ganey Associates, Inc.) and reports each physician's monthly percentile score for each facility worked. The EPMN then uses these PG scores to assess patient satisfaction.

RVUs/hr are currently used by the EPMN to measure physician productivity at each facility the physician works. However, the goal of the network management is not to simply compare physicians in a single facility, but how to motivate physicians to accept the assignments deemed necessary by the EPMN. Thus, both metrics must be modified to reflect each physician's performance within the entire network. Such metrics should objectively reflect a physician's relative standing in the EPMN, regardless of the number and location of the facilities at which the physician works or other facility characteristics. In the following, we discuss why the conventional simple averages of RVUs/hr and PG scores would be unfair to physicians working in multiple facilities and the need to propose an alternative.

4.1.1. Absolute Measure of Physician Performance. Let the total number of facilities physician i has worked be F_i. Within each facility j, the total number of patients visiting physician i is $V_{i,j}$ and the total number of hours physician i worked is $Hours_{i,j}$. The RVUs incurred by physician i's patient k in facility j is denoted as $RVU_{i,j,k}$. Thus, the total number of RVUs generated by physician i in facility j during the 54-month period are given by $TotalRVU_{i,j} = \sum_{k=1}^{V_{i,j}} RVU_{i,j,k}$. The RVUs/hr performance measure for physician i in facility j can be expressed as

$$RVUs/hr_{i,j} = \frac{Total\ RVUs_{i,j}}{Hours_{i,j}} \quad (5)$$

Patients/hr and RVUs/Patient can be computed for physician i in facility j similarly.

For a physician working at multiple facilities within the EPMN, RVUs/hr = (sum of RVUs in the network)/(sum of hours worked in the network); network Patients/hr and RVUs/Patient can be obtained similarly. Yet, these measures are facility-dependent. Facility demand affects service times, which directly influences the number of patients a physician cares for hourly (KC and Terwiesch 2009). Thus, comparing physicians across facilities using these metrics would be unfair, but RVUs/hr performance targets set by individual facilities are inadequate for evaluating EP productivity in network settings. Our data indicate the average EP within the EPMN works at 2.46 facilities, and EP groups often contract with health care systems with multiple EDs. These EDs have varying capabilities, patient catchment areas and staffing models. Moreover, EPs do not control patient arrivals and work various clinical schedules to cover facilities 24/7. The high degree of variability in patient volume and acuity leads to fluctuations in RVUs/hr, making the conventional approach invalid and necessitating more sophisticated methods to measure EP productivity.

4.1.2. Proposed Relative Indices for Physicians in Network. The new performance indices proposed here are better suited for concurrently evaluating all physicians working in a large EPMN, where a physician may be assigned to multiple facilities. Equation (5) quantifies physician i's performance in facility j. We define facility j's average performance as:

$$Average\ RVUs/hr_j = \frac{\sum_{i=1}^{N}\sum_{k=1}^{V_{i,j}} RVUs_{i,j,k}}{\sum_{i=1}^{N} Hours_{i,j}} \quad (6)$$

Taking the ratio Equation (5)/Equation (6), i.e., $\frac{RVUs/hr_{i,j}}{Average\ RVUs/hr_j}$, we know how the physician performs relative to peers in the specific facility, j. Next, we weight the ratio by the hours worked in each facility to derive a composite rating across all facilities (Equation (7)). Indexing this metric acknowledges the inability of EPs to control for demand, capacity, and competencies in the facilities worked.

$$RVUs/hr\ Index_i = \frac{\sum_{j=1}^{F_i}\left[\frac{RVUs/hr_{i,j}}{Average\ RVUs/hr_j} \times Hours_{i,j}\right]}{\sum_{j=1}^{F_i} Hours_{i,j}} \tag{7}$$

We define the volume index (Equation (8)) and the complexity index (Equation (9)) similarly, using hours in a facility and number of patients treated in a facility, respectively, as the weights:

$$Patients/hr\ Index_i = \frac{\sum_{j=1}^{F_i}\left[\frac{Patients/hr_{i,j}}{Average\ Patients/hr_j} \times Hours_{i,j}\right]}{\sum_{j=1}^{F_i} Hours_{i,j}} \tag{8}$$

$$RUVs/Patient\ Index_i = \frac{\sum_{j=1}^{F_i}\left[\frac{RUVs/Patient_{i,j}}{Average\ RUVs/Patients_j} \times V_{i,j}\right]}{\sum_{j=1}^{F_i} V_{i,j}} \tag{9}$$

The patient experience index (Equation (10)) is also computed using hours worked in a facility as the weights:

$$PG\ Index_i = \frac{\sum_{j=1}^{F_i}\left[\frac{PG\ Score_{i,j}}{Average\ PG\ Score_j} \times Hours_{i,j}\right]}{\sum_{j=1}^{F_i} Hours_{i,j}} \tag{10}$$

Note that an index of 100% indicates that a physician is performing on par with peers in the network; 110% implies 10% better performance than peers; while 85% suggests that performance is 15% inferior to peers.

Advantages of the proposed indices are discussed in a Technical Supplement S.3, where we compare Equations (7)–(9) with alternate metrics. Technical Supplement S.3 also details the mathematical rationale for supporting the indices we proposed.

4.1.3. The Need for the New Physician Network Performance Indices. Normalizing by the facility averages removes the effects of scale (facility average) and addresses the effects of relative hours worked at various facilities in the EPMN. Such a method is fair since facility assignment is often beyond physicians' control, given the need of the EPMN to staff all contracted facilities and the competitive nature of obtaining more health care facilities to contract with the EPMN. Table 2 gives a stylized example, which resembles information found in the data, to highlight the need and advantage of using the proposed indices (Equation (7)). Physician A worked in two facilities (Facilities 1 and 2), while Physician B only worked in Facility 2. Facility 1 has higher average RVUs/hr (14 vs. 4).

Since Physician B only works at one facility, calculating RVUs/hr is straightforward, but facility differences complicate the calculation for Physician A. First, we examine the absolute measures, which is the ratio of Total RVUs to Total Hours. In that case, Physician A [6.0 = (9000 + 3000)/(1200 + 800)] generates more RVUs/hr than Physician B [4.1 = 9000/2200] as shown in Table 2. Based on the absolute RVUs/hr measure, Physician A is more productive (6.0) than Physician B (4.1). We then compare physicians on the relative indices, detail their derivations below, and contrast this with the absolute measures.

We first calculate the ratios for each facility where the physician works. Physician A's RVUs/hr in Facility 1 is 7.5 (= 9000/1200), while Facility 1's average RVUs/hr is 14. Thus, Physician A is a below average performer in Facility 1, with a ratio of 0.54 (= 7.5/14) indicating that performance is 46% below peers. Physician A also appears to be slightly below average at Facility 2 (0.94). Using Equation (7), the index is:

$$RVUs/hr\ Index: 69.6\% = (0.54 \times 1200 + 0.94 \times 800)/(1200 + 800)$$

The indices in Table 2 show that Physician A is below average with regard to RVUs/hr (69.6%) and Physician B is above average with regard to RVUs/hr (102.3%). The absolute measure and the indexed measure give different conclusions about Physicians A and B's performances. By reflecting the relative performance at facilities in which a physician works and weighing performance on the time spent in each facility, the index proposed in Equation (7) neutralizes the scale (facility size) bias. Thus, the proposed index better reflects the true performance when the network

Table 2 The Need of Using Performance Indices – An Example

					RVUs/hr				
Phys.	Fac.	RVUs	Patients	Hours	Fac.	Phys.	RVUs/hr Ratio	RVUs/hr	RVUs/hr Index
A	1	9000	3600	1200	14.0	7.5	0.54	6.0	**69.6%**
	2	3000	900	800	4.0	3.8	0.94		
B	2	9000	2500	2200	4.0	4.1	1.02	4.1	**102.3%**

administrators assign facilities. This demonstrates the importance of adopting the relative index for assessing physician performance within large networks where physicians work at multiple facilities.

4.1.4. Comparing Network Physician Performance Using Proposed Index. We find a strong positive correlation between RVUs/hr Index and absolute RVUs/hr (r = 0.7687, p-value < 0.001), confirming that a substantial portion of revenue potential (RVUs/hr) can be explained by the proposed index. It captures the "intrinsic" ability of a physician to potentially generate revenue across facilities since it weights relative performance by hours at the various facilities. The unexplained portion may be driven by patient characteristics. Thus, the index is an equitable measure of physician performance.

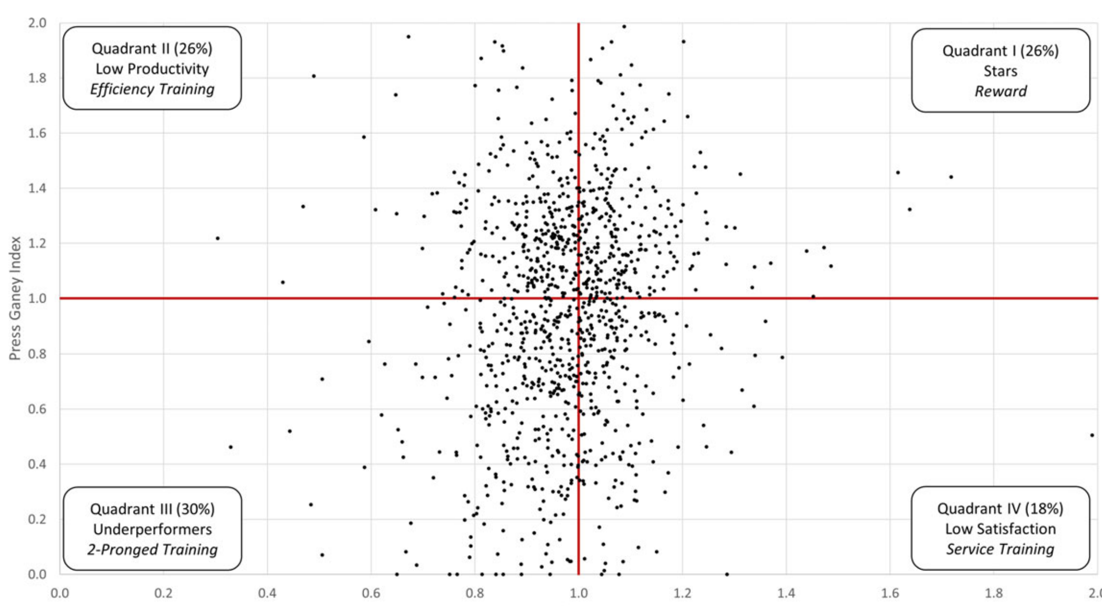
To provide a more comprehensive view and offer appropriate "carrots and sticks" for reward and remedy, we jointly consider both productivity and patient satisfaction by plotting the new RVUs/hr Index (x-axis) and PG Index (y-axis) on the xy-plane. Depending on which quadrant a physician falls, management can easily identify the need for improvement with regard to productivity, patient satisfaction, or both, or if the physician should be rewarded for outstanding performance on both dimensions (Figure 4).

Figure 4 displays physicians' performances, where each dot represents one physician. The horizontal and vertical lines signify average performance on each dimension, and thus divide physicians into four groups (quadrants). Management may choose cutoff values within each quadrant to further differentiate physicians within each group, e.g., rewarding those who perform one standard deviation above the mean (center point) on both dimensions. Additionally, they may choose to weight the two measures differently depending on specific performance goals. The "best" physicians (26%) who deliver above average performance in both productivity and patient satisfaction, i.e., RVUs/hr Index ≥ 1 and PG Index ≥ 1, are located in Quadrant I and may be used for benchmarking. Physicians who fall in Quadrant III (30%) are below average performers on productivity and patient experience, and require immediate attention from management. Quadrant II (26%) shows those short on productivity, while Quadrant IV (18%) depicts those with low patient experience; both II and IV are candidates for remedial training depending on the specific threshold set by management. It may be necessary to further address ways to balance patient experience and productivity in Quadrants II and IV.

5. Cluster Analysis to Manage Physician Segments

We have proposed a fairer, systematic, logical and comprehensible performance metric to objectively evaluate physicians working in an EPMN in section 4. To identify individual traits that may affect physician performance, we group physicians based on non-performance attributes and examine the disparities among clusters. Such information can be used to guide and reward different groups of physicians; and

Figure 4 **Integrating Productivity and Patient Experience: Insights for Rewards and Improvement Training [Color figure can be viewed at wileyonlinelibrary.com]**



more efficiently and effectively allocate physicians to facilities to meet the needs of the EPMN.

Due to known differences between pediatric and general EDs, we treat pediatric physicians as a separate group. Since pediatric physicians only account for 5% of the physicians in our data, we have excluded physicians from the cluster analysis if the average age of their patients was less than 18. Using the variables in Table 3, we group the 1026 non-pediatric EPs into five fundamentally different clusters using K-means.

We have colloquially named clusters based on dominant characteristics (e.g., Night Owl, Veteran, etc.) Figures 5a and b graphically display differences among the five clusters. The dashed line represents the overall average, while each boxplot shows the distribution of the variable on the *y*-axis. The plot provides the (1) minimum, (2) 1st quartile (25th percentile), (3) median (50th percentile), (4) 3rd quartile (75th percentile), and (5) maximum. Figures 5a and b shows that Veteran physicians are older and tend to work fewer nights, while Night Owls work a disproportionately high number of night shifts relative to others. Minority and Female physicians are generally younger, consistent with medical school trends.

By zooming in on Figure 4 and examining the five clusters separately, we can link the cluster membership with performance relative to the entire physician population. Figure 6a shows a relatively even distribution of Veterans across the four quadrants, while Figure 6b suggests that the majority of Night Owls have low RVUs/hr indices, with many also earning lower PG scores. One-way ANOVA has confirmed that the cluster differences in PG scores and RVUs/hr Index are statistically significant.

Table 3 Summary Statistics of Variables used in Cluster Analysis

	Average patient age ≥ 18 (N = 1026)	
Variable	Mean	SD
Physician Age	45.749	9.967
# Facilities Worked	2.518	2.334
% 6AM-3PM Hours	36.724	14.257
% 3PM-12PM Hours	47.139	11.342
% 12PM-6AM Hours	16.126	15.176
% Admitted	18.052	10.018
APP Support Ratio (%)	23.268	10.226
Variable	**Count**	**Percent**
Male	715	69.688
White	812	79.142

Thus far, we have provided a structural, evidence-based approach to segment EPs within the network. The clusters suggest that EPs may have different priorities at different stages in their personal and professional lives. For instance, younger physicians (i.e., many female and minority physicians) may have different family and financial concerns than veteran doctors (Darcy et al. 2012, Dyrbye et al. 2013). Like most professions, the composition of EPs encompasses varying career ambitions and personal responsibilities. It is thus vital to understand such distinctions when matching physicians with divergent priorities to different shifts and multiple facilities. Specifically, given the need to maintain continuous and universal availability of EDs, managers should consider physician characteristics when pursuing operational productivity and greater patient experience. There are legal reasons why we do not recommend linking compensations with these clusters. By law, employers

Figure 5 Distributions of Age and Night Shift by Clusters [Color figure can be viewed at wileyonlinelibrary.com]

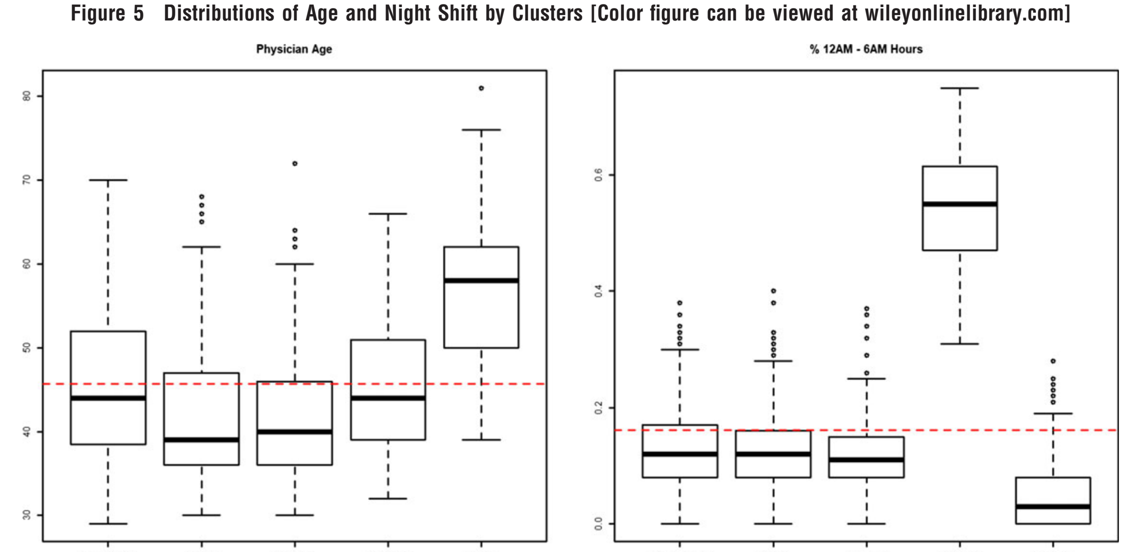
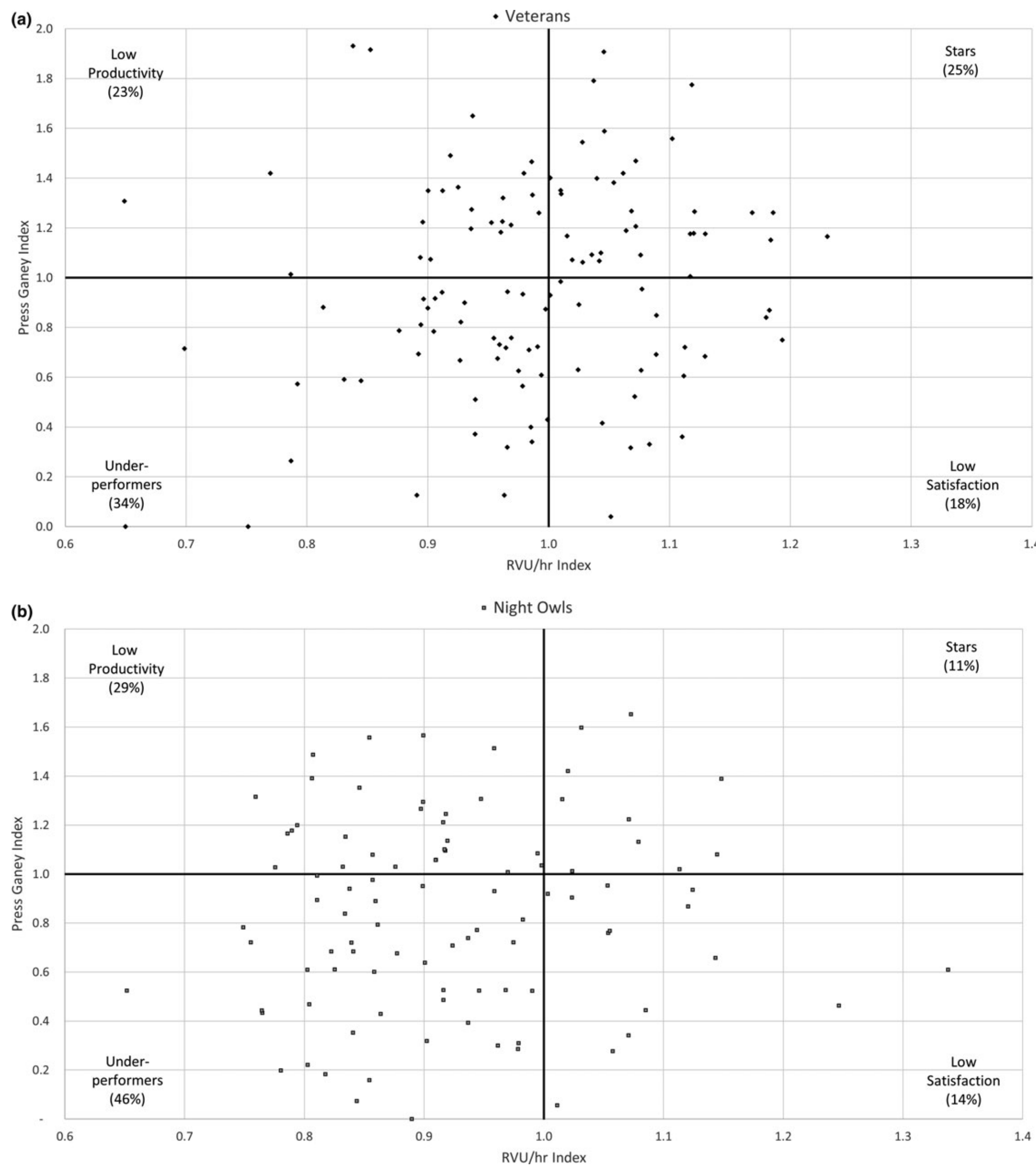


Figure 6 RVUs/hr Index (Productivity) vs. PG Index (Patient Experience) by Cluster (a) 25% of Veterans (Quadrant I) Achieve above Average Performance on both Dimensions (b) 75% (Quadrants II & III) of Night Owls Earn below Average RVUs/hr Index



cannot compensate based on age, gender, or race. However, the cluster analysis reveals that physicians are naturally grouped by these factors. Thus, these clusters only help to develop insights for our study, instead of serving as the basis for compensation.

The clustering results lead us to reflect on current practices. Namely, management must appreciate the dynamics of doctors' professional and personal development and recognize that physicians will respond distinctively to incentives. Measuring performance collectively for all clusters or offering identical incentives to all physicians may not be effective. For example, comparing Night Owls with Veterans who work fewer night shifts could be unfair due to differences in patient composition. In the following section, we use the cluster differences to motivate our selection of control variables.

6. Empirical Results

The indices in Equations (7)–(10) neutralize the impact of exogenous patient demand and allow for fairer comparisons among physicians operating in an EPMN. In this section, we first model RVUs/hr Index to demonstrate how the four indices are linked. Then, we identify the drivers of these indices, i.e., to understand how patient, physician, and facility factors impact physicians' relative indices. Finally, we analyze the relationship between the revenue potential index and PG patient experience index to discern if physicians with high revenue potential sacrifice patient experience scores.

6.1. Drivers of Relative Indices: A Simultaneous Equations Model

Through a log-log model (not presented here), we empirically show that the product of the volume and complexity indices explains 99.86% of the variability in RVUs/hr Index. Namely, ln(RVUs/hr Index) = 1.00∗ln(Patients/hr Index) + 0.99∗ln(RVUs/Patient Index). While this product does not necessarily result in the RVUs/hr Index mathematically, the strength of this empirical relationship led us to further examine the indices' relationships. Technical Supplement S.4 proves that the multiplicative equivalence holds when the average flow rates among facilities are equal. As the facility flow rates in our dataset are quite comparable for a given physician, this explains the strong empirical result. Specifically, the average facility flow rate deviation for a physician is 0.3, corresponding to a mean absolute percentage difference (MAPD) of 8.6%. Note that Physician i's MAPD is computed by

$$MAPD_i = \frac{\sum_{j=1}^{F_i} \frac{|Patients/hr_j - Average\ Patients/hr|}{Average\ Patients/hr}}{F_i},$$

with F_i denoting the total number of facilities in which physician i has worked.

The coefficients for the volume and complexity indices are both approximately equal to positive one and highly significant, and these effects are synergistic. As both indices strongly influence the RVUs/hr Index, physicians can enhance their RVUs/hr Index through improving volume or complexity performance (or both). Subsequently, we explore how exogenous factors affect complexity and volume indices, and thus indirectly drive revenue potential, using a system of simultaneous equations. In addition to the control variables identified through clustering, we chose to include physician characteristics and variables that change shift-by-shift, e.g., patient population, type of shift (day, afternoon, or night), etc. While geographic location, demographics around the facility, the number of beds, etc. are important factors, these are facility related variables, which are indirectly accounted for in our indices.

The first and second stage estimates from the two-stage least squares (2SLS) model (Equations (1)–(2)) are summarized in Tables 4a and b, respectively, and we have confirmed the robustness of these results using bootstrap standard errors (see Technical Supplement S.2). We utilize the 2SLS regression procedure to simultaneously estimate the two equations because it accounts for correlations between endogenous variables (volume and complexity indices); and between endogenous variables and the 2nd-stage errors. We confirm that this system is identified based on the rank condition, which is a necessary and sufficient condition for identification. More specifically, each of the control variables excluded from one equation appears in the other equation. Equation (1) in Table 4b corresponds to the complexity index and comprises patient characteristics (e.g., Patient Age, % Male Patients, % Admitted, and ICD-9 Code groups). These controls are not included in Equation (2) for the volume index because physicians do not schedule or select their patients. In contrast, physician characteristics (e.g., age, gender, and race) and coding communications per patient are only included in Equation (2). Other exogenous variables (insurance types, shifts worked, and the APP support ratio) overlap between the two equations, signifying both their direct and indirect effects on the volume and complexity indices.

The coefficient for ln(Patients/hr Index) in Equation (1) is significant and positive (β = 0.1841), which is inconsistent with our first hypothesis (H1). Conversely, the coefficient for ln(RVUs/Patient Index) in Equation (2) is significant and negative (β = −0.8224), suggesting an inverse relationship and providing support for H1. Thus, the simultaneous equations model partially supports H1. Table 4a and b indicates that many factors influence the volume and complexity indices, and the relationships between the exogenous variables and the indices involve both direct and indirect effects. This demonstrates the extent to which the volume and complexity indices are intertwined with each other and with exogenous factors.

We have thus far analyzed the factors influencing the volume and complexity indices and how these two indices drive RVUs/hr Index, but does the need to increase physician productivity in the EPMN result in reduced patient experience? We next examine how revenue potential (efficiency) affects patient experience.

6.2. Drivers of PG Score

CMS (2015) has recently introduced Value-based Purchasing (VBP) programs. These quality initiatives financially motivate hospitals to improve upon current operations, as reimbursement now depends on patient outcomes and patient experience. The PG survey is a widely employed tool for gauging perceptions among discharged patients. The survey includes questions related to the individual physician and the facility, and both physician and facility PG scores (percentiles) are reported monthly. In this section, we

Table 4 2SLS Model: Relationship between Volume and Complexity Indices (*N* = 1079)

		Y = ln(RVUs/Patient Index)		*Y* = ln(Patients/hr Index)	
Level	Variable	Coefficient	SE	Coefficient	SE
(a) First-stage Regression					
	(Constant)	−0.4735	0.1040	0.4202	0.3025
Physician	ln(RVUs/Patient Index)				
	ln(Patients/hr Index)				
	Peds Indicator (0 = General, 1 = Peds)	0.4418***	0.0610	−0.4752**	0.1775
	Physician Age	−0.0001	0.0004	0.0008	0.0012
	Physician Male (0 = Female, 1 = Male)	−0.0022	0.0038	0.0365**	0.0112
	Physician White (0 = Non-white, 1 = White)	0.0031	0.0042	0.0303*	0.0123
	% 12AM - 6AM Hours	0.0701***	0.0180	−0.1119*	0.0523
	% 6AM - 3PM Hours	−0.0322	0.0176	0.2241***	0.0512
	Coding Com/Patient	−0.0097**	0.0037	−0.0402***	0.0107
	APP Support Ratio	0.1468	0.0748	0.4859*	0.2174
	APP Support Ratio*Physician Age	−0.0025	0.0016	−0.0090*	0.0046
Patient	Average Patient Age	0.0289***	0.0039	−0.0257*	0.0113
	(Average Patient Age)^2	−0.0003***	0.0000	0.0003*	0.0001
	% Male Patients	−0.0232	0.0863	0.1888	0.2510
	Commercial Index	0.0032	0.0294	−0.2592**	0.0855
	Medicaid Index	−0.1902***	0.0129	0.1274**	0.0375
	Medicare Index	0.0371***	0.0066	−0.1053***	0.0193
	Self-Pay Index	−0.2229***	0.0202	0.3605***	0.0589
	% ICD9 Group 1	0.3313***	0.0394	−0.2913*	0.1147
	% ICD9 Group 2	0.2155***	0.0306	−0.1530	0.0890
	% ICD9 Group 3	0.0582	0.0495	−0.2163	0.1440
	% Admitted Patients	0.1255***	0.0233	−0.1552*	0.0678
Adjusted R^2		64.64%		27.51%	

		Equation (1): ln(RVUs/Patient Index)		Equation (2): ln(Patients/hr Index)	
Level	Variable	Coefficient	SE	Coefficient	SE
(b) Second-stage Regression					
	(Constant)	−0.5648	0.1258	0.0351	0.1626
Physician	ln(RVUs/Patient Index)			−0.8224***	0.1391
	ln(Patients/hr Index)	0.1841**	0.0659		
	Peds Indicator (0 = General, 1 = Peds)	0.5384***	0.0809		
	Physician Age			0.0002	0.0012
	Physician Male (0 = Female, 1 = Male)			0.0367**	0.0109
	Physician White (0 = Non-white, 1 = White)			0.0337**	0.0120
	% 12AM - 6AM Hours	0.0839***	0.0228	−0.0191	0.0504
	% 6AM - 3PM Hours	−0.0811**	0.0268	0.2205***	0.0488
	Coding Com/Patient			−0.0457***	0.0104
	APP Support Ratio	0.0170	0.0232	0.5794**	0.2132
	APP Support Ratio*Physician Age			−0.0098*	0.0044
Patient	Average Patient Age	0.0345***	0.0050		
	(Average Patient Age)^2	−0.0004***	0.0001		
	% Male Patients	−0.0768	0.1064		
	Commercial Index	0.0476	0.0395	−0.2756**	0.0824
	Medicaid Index	−0.2124***	0.0183	−0.0485	0.0497
	Medicare Index	0.0564***	0.0107	−0.0650**	0.0202
	Self-Pay Index	−0.2929***	0.0357	0.1876**	0.0720
	% ICD9 Group 1	0.3697***	0.0519		
	% ICD9 Group 2	0.2598***	0.0368		
	% ICD9 Group 3	0.0865	0.0598		
	% Admitted Patients	0.1629***	0.0302		

Note: Independent variables defined in Tables A1 and A2.

examine the linkage between the proposed revenue potential index and patient experience scores. Specifically, we consider whether physicians trade off the RVUs/hr Index for the PG Index. For this, we employ OLS regression and model the physician PG Index as a function of the RVUs/hr Index and exogenous control variables (Table 5). The specific variables chosen for the model are based on the recommendations of

Table 5 PG Index Model Results

Dependent variable: PG Index†

Level	Variable	Coefficient	SE
	(Constant)	0.8027	0.2200
Physician	RVUs/hr Index	0.3234***	0.0719
	Physician Age	−0.0062***	0.0014
	Physician Male (0 = Female, 1 = Male)	0.0480	0.0291
	Physician White (0 = Non-white, 1 = White)	0.1421***	0.0323
	% 12AM–6AM Hours	−0.5652***	0.1027
	% 3PM–12AM Hours	−0.3506**	0.1308
	Efficiency Flag (0 = Not Complete, 1 = Complete)	0.0651*	0.0303
	Patient Satisfaction Flag (0 = Not Complete, 1 = Complete)	−0.0772**	0.0292
Patient	Average Patient Age	−0.0052***	0.0015
	% ICD9 Group 1	−0.1543	0.2867
	% ICD9 Group 2	0.7095**	0.2150
	% ICD9 Group 3	1.1821***	0.3447
	R^2	10.87%	
	Adjusted R^2	9.85%	

Notes: $***p < 0.001$, $**p < 0.01$, $*p < 0.05$; two-tailed tests. Independent variables defined in Tables A1 and A2.
†Refer to Equation (10) to derive the PG Index.

the clinicians and administrative staff of the EPMN under study.

The results in Table 5 show that almost 10% of the variability in the PG Index is explained by our model, and most predictors are statistically significant. Specifically, we observe a positive and significant coefficient for RVUs/hr Index (β = 0.3234) that supports our second hypothesis (H2).

Other significant factors in Table 5 suggest that PG Index is further influenced by physician-specific characteristics, such as work schedules as well as the average patient age. For example, physicians who work more night shifts receive significantly lower scores relative to their peers, possibly due to the acuteness of patients' conditions, stress experienced during night visits, and lower staffing levels at night. Working more hours between 3 pm and 12 am is also associated with lower relative scores. Furthermore, physicians seeing older patients receive lower scores relative to their peers on average. As physicians with lower PG scores are often required to complete the Patient Satisfaction training, we find a negative relationship between training completion and PG Index. Due to this bias and because training completion dates are not reported, we must be cautious when interpreting such results. We recommend that future data collection in the EPMN include the dates and reasons for training.

Table 5 suggests that a significant relationship exists between the RVUs/hr Index and PG Index. However, we speculate that unobserved factors play important roles in explaining relative PG scores and hence the relatively low explanatory power of the model.

The results of these hypotheses are important because this is how physicians get paid and how they are evaluated. It also helps to have objective metrics to compare health care providers so management can make better provider evaluation and incentive decisions. The primary driver of RVUs/hour is better training in multi-tasking to manage multiple patients. The primary driver of PG scores is to reduce length of stay for discharged patients (see Pines et al. 2018), and there are clearly proven techniques for improving communication to increase PG. Thus, the proposed metrics allow for such evaluations while taking into account factors largely outside of an individual physician's control. The indices provide insight into what is fixed and what needs to be adjusted.

7. Managerial Insights, Limitations, and Conclusions

7.1. Managerial Insights

7.1.1. Benchmarking Physician Performance and Continuous Improvement. Our findings have practical implications for EDs. Indexing physician metrics such as Patients/hr, RVUs/Patient, RVUs/hr, and PG Scores to facility-based averages mitigates the exogenous factors that affect physician performance. The proposed indices provide simple and intuitive benchmarks for objective evaluation of physicians. These indices facilitate physician segmentation into high and low performers, which initiates new management processes to incentivize and train physicians. Management may opt to customize the evaluation and ranking process by providing weights for each dimension of performance. Moreover, resulting clusters highlight distinct differences among physicians and provide further managerial insights.

Health care administrators are constantly looking for best practices to benchmark performance and improve processes. EDs are no exception, as evidenced by the Emergency Department Benchmarking Alliance (Wiler et al. 2015). However, benchmarking practices typically occur at the facility-level. With our study, benchmarking at the physician level both within a facility and across facilities is possible. That is, by identifying characteristics of the highest performing physicians in terms of RVUs/hr Index, other physicians can recognize factors under their control and strive for first-rate performance. Furthermore, the proposed indices provide objective measures of the network performance of physicians given that demand is beyond their control. The new indices capture physician performance relative to their peers and are highly correlated with the revenue potential of the physician in the EPMN. These indices neutralize the exogenous demand effects while capturing relative effort compared to peers. From the perspective of the administrators, the index approach is fair, simple, intuitive, and effective.

Our statistical models acknowledge that the same productivity level may not be attainable for all physicians at all facilities. We recommend that managers incorporate relative indices and provide adjustment factors when evaluating physicians. Exogenous demand and other facility-level factors impact the indices, and thus our 2SLS regression model (Table 4) can be used to adjust for various physician-specific and patient-specific factors to support more equitable comparisons of physicians' RVUs/hr performance. At a minimum, our facility-adjusted indices reduce the exogenous variability inherent to the respective absolute measures, and ED management could use our model to set performance standards. These results are helpful in tracking continuous improvement of physicians over time as well. Recognizing these results could lead to the development of new training programs or instituting mandatory completion of existing training for certain physicians. The clustering results reinforce our case for defining performance objectives (and incentives) pertaining to distinct physician segments. The clusters provide guidance for management to set physician-specific targets, schedule and allocate EPs, and implement differential physician "care and recruitment" strategies.

7.1.2. Management and Physician Benefits. While EDs typically cannot control the type or number of patients arriving for treatment during a given shift, they should recognize that patient experience scores are generally lower for physicians working mostly night shifts. In addition, physicians with more APP support tend to achieve higher volume indices, demonstrating that they treat more patients per hour. The proposed indexing approach is valuable for accurate assessment of physician productivity and could lead to more equitable compensation across physicians as the proposed performance indices are generalizable and applicable to various compensation models, such as fee-for-service, capitated, or value-based reimbursement. While physicians may favor daytime shifts for physiological reasons, our results indicate that performance differs significantly between day and night shifts. Physicians who work more night hours are somewhat penalized in both PG scores and RVUs/hr. Thus, management should take note of physicians willing to work nights in their facility assignment and in their compensation decisions as our empirical results provide compelling evidence and rationale for differentiating recompense. This information may even translate into scheduling decisions. For instance, a facility might stipulate a minimum proportion of night hours for all physicians in order to address the night shift effect. As physicians age, working nights becomes physically more demanding; compensation adjustments may need to follow to allow for fewer night shifts.

EPs may also gain from understanding our results. Since RVUs/hr is used to evaluate the physicians, they can benefit from recognizing the factors that affect productivity, especially those factors within their control (e.g., shift preference). This knowledge may help physicians to improve their productivity by controlling its drivers. It may also aid them in employment decisions. Physician involvement in shift choices offers them the ability to take into account both personal and financial concerns to make educated decisions that best fit their needs. All stakeholders benefit from physician productivity improvements, and understanding and tracking productivity is a critical step toward improving ED operations.

We have had discussions with the organization in question on piloting the use of these indices in particular health systems to prospectively assess their values. Currently, EPMN management employs arbitrary benchmarking that does not adjust for facility capabilities, physicians' career stages, and patient differences across EDs and shifts. Testing and subsequent implementation of these transparent and easy to implement indices with follow-up assessment is our planned next step.

7.2. Limitations

We have taken into account several control variables in our study of EPs, but other factors could also affect physician performance. Capacity variables such as the number of ED beds, nursing support, and the availability of diagnostic tools in a facility can influence processing rates as well. One limitation of our study is the inability to account for these effects. Moreover, we did not consider the impact of information technology. For example, some facilities have recently implemented electronic health record systems, which temporarily slow down operations and negatively impact ED performance (Ward et al. 2014a,b). While we examined productivity, providers and hospitals use several other clinical metrics to assess physician performance, such as patient outcomes and benchmark goals put forth by national organizations such as CMS. We focused on the operational performance metrics and have yet to link them to these other markers.

From a modeling perspective, we aggregated data over the entire study period (54 months) to analyze overall EP performance. However, this data may be modeled at a monthly level. Performing a cross-sectional analysis limits our ability to explore learning effects over time. Studying the data longitudinally would allow us to examine the influence of exogenous shocks on demand. For instance, outbreaks of Zika virus, Ebola virus, and various influenzas impact ED demand in prominent facilities, but we do not view the effects of these exogenous shocks when data is aggregated over many months. Future research could study EP performance on a more granular level, including quasi-experiments to expose causal relationships.

Furthermore, the way in which PG scores are reported raises some concerns about their usage. A physician's scores at any time are based on the PG surveys completed by previously seen patients at a given facility. Therefore, PG scores are not reported until a number of patients from that facility submit their surveys (possibly requiring several weeks' worth of data to reach an adequate sample). As with any survey, there may also be response bias, and PG percentile scores could fluctuate significantly from month to month.

Finally, implementation of the proposed indices for benchmarking may not be easy. Our discussions with physicians indicate some opposition to management by numbers alone. Thus, physician resistance, actionable items resulting from benchmarking, and understanding the meaning of these indices by a health care audience will likely present some challenges.

7.3. Conclusions

Our study is unique in several ways. First, we use big data to conduct a multi-facility, multi-year study of physician performance within a large EPMN, which is highly relevant as these physician networks continue to expand, especially in emergency medicine. We develop performance metrics that adjust for facility-level differences and allow for objectively comparing physicians who work at multiple facilities across a large and diverse EPMN. Our proposed indices overcome the deficiencies of existing physician performance measures that are only appropriate in the single-facility case and make equitable comparisons of physicians within the network possible.

We empirically demonstrate the value of the proposed indices in evaluating physicians within this large network. We subsequently use cluster analysis to identify physician segments with similar characteristics, which helps management to recognize physicians' priorities and to better understand which factors are driving physician performance. We verify that the proposed volume and complexity indices explain a substantial portion of variation in revenue potential productivity across physicians. The 2SLS model (Table 4) simultaneously examines the linkage between the volume and complexity indices and their drivers. Finally, we explore the relationship between physician productivity and patient experience (Table 5).

The 24/7 nature of EDs, the high incidence of burnout and the overall shortage of EPs relative to patient need make it necessary to effectively manage EPs across their entire careers. This research is highly relevant to the EM field as large physician management groups attempt to balance professional career satisfaction and the needs of large health systems for ED services. As consolidation occurs throughout health care in general and EM in particular, more physician management networks are emerging. The need for more robust data-driven measurements of network physicians persists, as these performance metrics are crucial for maintaining market position and bargaining power. Refined tools to measure and evaluate productivity carry lessons for assessing and managing the operational efficiency of the health care system where the ED stands at an important nexus.

While we have focused on ED-specific issues and problems associated with management of large EPMNs, the research framework in Figure 2 may be adapted for application to other settings. For example, regional managers of retail stores or restaurant chains overseeing various facilities may face disparities in performance across different sites due to site-specific attributes. By identifying dimensions on which performance will be judged and indexing those measures relative to site averages, organizations can develop an equitable method to assess an employee's relative performance across sites. Thus, the proposed research framework (Figure 2), comprising the indexing system, performance matrix (quadrants), clustering, and driver identification, is generalizable and can be broadly applied after making industry-specific or company-specific adaptations. Our model is thus valuable for performance enhancement and employee development in data-intensive business settings.

Acknowledgments

This research was partially supported by the Center for Health Care Management at the University of Pittsburgh. The authors acknowledge US Acute Care Solutions for their help in compiling the data used in this study and thank Dominic Bagnoli, MD, Michael Osmundson, MD, James Augustine, MD, Amer Aldeen, MD, and the leadership of US Acute Care Solutions for their collaboration. The authors also acknowledge the critical review of the manuscript by Jestin N. Carlson, MD, Michael J. Ward, MD, MBA and Jesse M. Pines, MD, MBA, MSCE and the support of this project by Christopher Corbit, MD, Suzi Fix, JD, Paul Dietzen, and Jesse Eterovich, BS. Finally, the authors are grateful to the Editor and two anonymous reviewers for their constructive comments.

Appendix. Additional Tables and Variable Definitions

Similar to Equations (7)–(10), we propose Equation (A1) to index the proportion of insurance types handled by each physician. It reflects whether the physician is treating more or less patients of a certain insurance type than peers. Additionally, physician i's APP Support Ratio is the number of combined APP hours during all shifts divided by the total number of physician and APP hours during these shifts.

$$Insurance\ m\ Index_i = \frac{\sum_{j=1}^{F_i}\left[\frac{\%Insurance\ Type\ m_{i,j}}{Average\ \%Insurance\ Type\ m_j} \times V_{i,j}\right]}{\sum_{j=1}^{F_i} V_{i,j}} \tag{A1}$$

Table A1 Definitions of Patient Level Variables

Variable	Definition
Average Patient Age$_i$	= Average age of patients treated by a physician across all visits in all facilities
Peds Indicator$_i$	$= \begin{cases} 1, \text{if } Average\ Patient\ Age_k < 18 \text{ years} \\ 0, otherwise \end{cases}$
Coding Com/Patient$_i$	= Average number of communications with physician per patient visit due to unclear or potentially missing physician documentation
Admit Ratio$_i$	= Proportion of a physician's patients that were admitted
Commercial Insurance Ratio$_i$	= Proportion of a physician's patients with commercial insurance
Medicare Ratio$_i$	= Proportion of a physician's patients with Medicare
Medicaid Ratio$_i$	= Proportion of a physician's patients with Medicaid
Self-Pay Ratio$_i$	= Proportion of a physician's patients without insurance
Male Ratio$_i$	= Proportion of male patients
ICD-9 Group 1 Ratio$_i$	= Proportion of Circulatory, Respiratory, Digestive, and Genitourinary visits (390-629)
ICD-9 Group 2 Ratio$_i$	= Proportion of visits due to Symptoms, Signs & Ill-defined Conditions (780-799)
ICD-9 Group 3 Ratio$_i$	= Proportion of Injury & Poisoning visits (800-999)

Table A2 Definitions of Physician Level Variables

Variable	Definition
# Facilities Worked$_i$	= Number of facilities in which a physician worked
Physician Age$_i$	= Age of Physician
Physician Gender$_i$	$= \begin{cases} 1, \text{if } Male \\ 0, otherwise \end{cases}$
Physician Race$_i$	$= \begin{cases} 1, \text{if } White \\ 0, otherwise \end{cases}$
Efficiency Flag$_i$	$= \begin{cases} 1, \text{if } EPMN - administered\ Efficiency\ Training\ Completed \\ 0, otherwise \end{cases}$
Satisfaction Flag$_i$	$= \begin{cases} 1, \text{if EPMN} - administered\ Patient\ Satisfaction\ Training\ Completed \\ 0, otherwise \end{cases}$
6am-3pm Ratio$_i$	= Proportion of a physician's hours worked between 6 am and 3 pm
3pm-12am Ratio$_i$	= Proportion of a physician's hours worked between 3 pm and 12 am
12am-6am Ratio$_i$	= Proportion of a physician's hours worked between 12 am and 6 am
APP Support Ratio$_i$	= Proportion of provider hours worked during a physician's shifts which are concomitantly supported by APPs

References

ACEP. 2016. *CEDR Frequently Asked Questions*. Available at https://www.acep.org/globalassets/cedr_pdfs/CEDR-FAQs.pdf (accessed date August 27, 2017).

Angst, C. M., S. Devaraj, C. C. Queenan, B. Greenwood. 2011. Performance effects related to the sequence of integration of healthcare technologies. *Prod. Oper. Manag.* **20**(3): 319–333.

Armony, M., A. R. Ward. 2010. Fair dynamic routing in large-scale heterogeneous-server systems. *Oper. Res.* **58**(3): 624–637.

Ba, S., B. R. Nault. 2017. Emergent themes in the interface between economics of information systems and management of technology. *Prod. Oper. Manag.* **26**(4): 652–666.

Bhargava, H. K., A. N. Mishra. 2014. Electronic medical records and physician productivity: Evidence from panel data analysis. *Management Sci.* **60**(10): 2543–2562.

Blank, J. L., E. Eggink. 2014. The impact of policy on hospital productivity: A time series analysis of Dutch hospitals. *Health Care Manag. Sci.* **17**(2): 139–149.

Boudreaux, E. D., E. L. O'Hea. 2004. Patient satisfaction in the emergency department: A review of the literature and implications for practice. *J. Emerg. Med.* **26**(1): 13–26.

Boudreaux, E. D., B. L. Cruz, B. M. Baumann. 2006. The use of performance improvement methods to enhance emergency department patient satisfaction in the United States: A critical review of the literature and suggestions for future research. *Acad. Emerg. Med.* **13**(7): 795–802.

Brennan, D. F., S. Silvestri, J. Y. Sun, L. Papa. 2007. Progression of EM resident productivity. *Acad. Emerg. Med.* **14**(9): 790–794.

Büchner, V. A., V. Hinz, J. Schreyögg. 2016. Health systems: Changes in hospital efficiency and profitability. *Health Care Manag. Sci.* **19**(2): 130–143.

Clinkscales, J. D., F. M. Fesmire, J. R. Hennings, H. W. Severance, D. C. Seaberg, N. Patil. 2016. The effect of EM residents on clinical efficiency and staffing requirements. *Acad. Emerg. Med.* **23**(1): 78–82.

CMS. 2015. Hospital value-based purchasing. Available at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/index.html (accessed date September 5, 2016).

Darcy, C., A. McCarthy, J. Hill, G. Grady. 2012. Work–life balance: One size fits all? An exploratory analysis of the differential effects of career stage *Eur. Manag. J.*, **30**(2): 111–120.

Dyrbye, L. N., P. Varkey, S. L. Boone, D. V. Satele, J. A. Sloan, T. D. Shanafelt. 2013. Physician satisfaction and burnout at different career stages. In Mayo Clinic Proceedings (Vol. 88, No. 12, pp. 1358–1367). Elsevier.

Eitel, D. R., S. E. Rudkin, M. A. Malvehy, J. P. Killeen, J. M. Pines. 2010. Improving service quality by understanding emergency department flow: A White Paper and position statement prepared for the American Academy of Emergency Medicine. *J. Emerg. Med.* **38**(1): 70–79.

Ellaway, R. H., M. V. Pusic, R. M. Galbraith, T. Cameron. 2014. Developing the role of big data and analytics in health professional education. *Med. Teach.* **36**(3): 216–222.

Grönroos, C., K. Ojasalo. 2004. Service productivity: Towards a conceptualization of the transformation of inputs into economic results in services. *J. Bus. Res.* **57**(4): 414–423.

Handel, D. A., L. K. French, J. Nichol, J. Momberger, R. Fu. 2014. Associations between patient and emergency department operational characteristics and patient satisfaction scores in an adult population. *Ann. Emerg. Med.* **64**(6): 604–608.

Hwang, C. E., G. S. Lipman, M. Kane. 2015. Effect of an emergency department fast track on Press-Ganey patient satisfaction scores. *Western J. of Emerg. Med.* **16**(1): 34.

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

Kayyali, B., D. Knott, S. Van Kuiken. 2013. The big-data revolution in US health care: Accelerating value and innovation. *McKinsey and Company* **2**(8): 1–13.

KC, D. S. 2013. Does multitasking improve performance? Evidence from the emergency department. *Manuf. Serv. Oper. Manag.* **16**(2): 168–183.

KC, D. S., C. Terwiesch. 2009. Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Sci.* **55**(9): 1486–1498.

Lee, E. K., H. Y. Atallah, M. D. Wright, E. T. Post, C. Thomas IV, D. T. Wu, L. L. Haley Jr. 2015. Transforming hospital emergency department workflow and patient care. *Interfaces* **45**(1): 58–82.

Moskowitz, A., J. McSparron, D. J. Stone, L. A. Celi. 2015. Preparing a new generation of clinicians for the era of big data. *Harv. Med. Student Rev.* **2**(1): 24.

Murdoch, T. B., A. S. Detsky. 2013. The inevitable application of big data to health care. *JAMA* **309**(13): 1351–1352.

Nambiar, R., R. Bhardwaj, A. Sethi, R. Vargheese. 2013. A look at challenges and opportunities of big data analytics in healthcare. In Big Data, 2013 IEEE International Conference on (pp. 17–22). IEEE.

Obermeyer, Z., E. J. Emanuel. 2016. Predicting the future—big data, machine learning, and clinical medicine. *N. Engl. J. Med.* **375**(13): 1216.

Pines, J. M., S. Iyer, M. Disbot, J. E. Hollander, F. S. Shofer, E. M. Datner. 2008. The effect of emergency department crowding on patient satisfaction for admitted patients. *Acad. Emerg. Med.* **15**(9): 825–831.

Pines, J. M., P. Penninti, S. Alfaraj, J. N. Carlson, O. Colfer, C. K. Corbit, A. Venkat. 2018. Measurement under the microscope: High variability and limited construct validity in emergency department patient-experience scores. *Ann. Emerg. Med.* **71**(5): 545–554.e6. https://doi.org/10.1016/j.annemergmed.2017.11.011.

Powell, A., S. Savin, N. Savva. 2012. Physician workload and hospital reimbursement: Overworked physicians generate less revenue per patient. *Manuf. Serv. Oper. Manag.* **14**(4): 512–528.

Raghupathi, W., V. Raghupathi. 2014. Big data analytics in healthcare: Promise and potential. *Health Inf. Sci. Syst.* **2**(1): 3.

Reiter, M., L. S. Wen, B. W. Allen. 2016. The Emergency Medicine Workforce: Profile and projections. *J. Emerg. Med.* **50**(4): 690–693.

Song, H., A. L. Tucker, K. L. Murrell. 2015. The diseconomies of queue pooling: An empirical investigation of emergency department length of stay. *Management Sci.* **61**(12): 3032–3053.

Sonnati, R. 2015. Improving healthcare using big data analytics. *Int. J. Scienti. Technol. Res.* **4**(8): 142–146.

Srinivasan, U., B. Arunasalam. 2013. Leveraging big data analytics to reduce healthcare costs. *IT Profess.* **15**(6): 21–28.

Theokary, C., Z. J. Ren. 2011. An empirical study of the relations between hospital volume, teaching status, and service quality. *Prod. Oper. Manag.* **20**(3): 303–318.

Venkat, A., S. Kekre, G. G. Hegde, J. Shang, T. P. Campbell. 2015. Strategic management of operations in the emergency department. *Prod. Oper. Manag.* **24**(11): 1706–1723.

Ward, A. R., M. Armony. 2013. Blind fair routing in large-scale service systems with heterogeneous customers and servers. *Oper. Res.* **61**(1): 228–243.

Ward, M. J., C. M. Froehle, K. W. Hart, S. P. Collins, C. J. Lindsell. 2014a. Transient and sustained changes in operational performance, patient evaluation, and medication administration during electronic health record implementation in the emergency department. *Ann. Emerg. Med.* **63**(3): 320–328.

Ward, M. J., A. B. Landman, K. Case, J. Berthelot, R. L. Pilgrim, J. M. Pines. 2014b. The effect of electronic health record implementation on community emergency department operational measures of performance. *Ann. Emerg. Med.* **63**(6): 723–730.

Wiler, J. L., S. Welch, J. Pines, J. Schuur, N. Jouriles, S. Stone-Griffith. 2015. Emergency department performance measures updates: Proceedings of the 2014 Emergency Department Benchmarking Alliance Consensus Summit. *Acad. Emerg. Med.* **22**(5): 542–553.

Supporting Information

Additional supporting information may be found online in the supporting information tab for this article:

Appendix S1: Technical Supplement.