Outside the Hospital Walls: Associations of Value Based Care Metrics and Community Health Factors

Catherine Markley
Dept. of Computer Science & iCeNSA
University of Notre Dame, USA
cmarkley@nd.edu

Keith Feldman
Dept. of Computer Science & iCeNSA
University of Notre Dame, USA
kfeldman@nd.edu

Nitesh V. Chawla*
Dept. of Computer Science & iCeNSA
University of Notre Dame, USA
nchawla@nd.edu

***Abstract*—As the healthcare industry shifts from traditional fee-for-service payment to value-based care models, the need to accurately quantify and compare the performance of institutions has become an integral component of both policy and research. To date, several notable metrics have been introduced, including the Centers for Medicare and Medicaids Hospital Value Based Purchasing (HVBP) program. However, despite widespread adoption, these standards suffer from a fundamental oversight. Where the factors utilized to characterize performance reflect only intrinsic facets of an institutions care, capturing elements of mortality rates, patient satisfaction, outcomes, and spending. Yet, this approach is directly at odds with our current understanding of health and wellness, as it is well known that social, economic, and community factors are deeply intertwined with healthcare outcomes. To this end, with institutions spread across diverse geographic regions, our manuscript demonstrates that HVBP performance metrics do not exist in isolation. Rather, they possess strong associations to the community factors in which the institution resides. Aggregating a broad set of factors from disparate data sources, this work moves through the informatics pipeline. Identifying performance scoring profiles though clustering and employing robust linear models to uncover novel relationships and advance the discussion around the need for value-based care quality metrics.**

I. Introduction

Amidst an increasingly aged and chronically ill population, skyrocketing healthcare costs have exposed the dire need to improve the value of resources devoted to the U.S healthcare system [1]. Driven by initiatives such as the Institute for Healthcare Improvement Triple Aim and advancements in electronic records, the healthcare industry has undergone significant change to address such need; fostering shifts from reactive to preventative care and advancing personalized medicine [2]. Yet, perhaps, no change has impacted hospital policy and practitioner life to the degree of transitioning from fee-for-service reimbursement models to value-based care.

Under the historical fee-for-service paradigm, health systems received payment based on the volume of services provided, wherein a predetermined amount was paid for each service regardless of outcome [3]. However, in line with efforts to improve population health and drive down costs, value-based care systems have emerged offering payment based on the quality of care provided. While numerous approaches to value-based care have been proposed, among the most successful has been the emergence of value-based purchasing frameworks; where "providers are paid fee-for-service with payment adjustments up or down based on value metrics" [4].

These frameworks reached a significant milestone when in 2013 the Centers for Medicare and Medicaid Services (CMS) introduced the Hospital Value-Based Purchasing (HVBP) program [5]. Implemented in acute care hospitals across the U.S., the program quantifies and compares hospital performance through an array of quality measures. These comparisons are in turn tied to financial incentives, today accounting for 2% of all Medicare payments to participating hospitals [6].

As such, it is unsurprising that HVBP metrics have garnered a great deal of attention from hospital administers and their employees. Although much of the focus has been on optimizing performance to improve rankings, a deeper consideration of the metrics themselves reveals an interesting facet of their computation. Representing four primary scoring domains, the metrics are comprised of numerous subcategories capturing inwardly looking elements of hospital care, such as mortality rates, patient satisfaction, outcomes, and spending.

However, with hospitals embedded in diverse communities across the country, such an approach overlooks established ties of economic and social factors to health outcomes. To date, little is understood about the relation between such community factors and HVBP performance. Yet, with the comparison between hospitals acting an integral aspect of the value-based purchasing payment model, the influence of such factors represents a potential source of unaddressed bias.

Our work addresses these relations, identifying associations between the HVBP performance metrics and a wide array of socioeconomic and community health factors. Ultimately, presenting a strong case for the need to broaden our current views, and account for external factors of a hospital's community.

In addressing this novel question, our manuscript moves through the complete informatics pipeline. First, extracting community factors from disparate government and academic datasets, and associating them with the HVBP scores of hospitals in their geographic region. Next, utilizing a clustering approach to account for the interconnected nature of HVBP performance metrics to identify scoring profiles. Finally, employing linear models to identify significant relations between these profiles and factors of the hospital's community.

*Corresponding Author
This work is supported in part by the NSF Grant IIS-1447795.

978-1-7281-0848-3/19/$31.00 ©2019 IEEE

II. Related Works

The notion of evaluating hospital performance has a long history, originating well before the emergence of VBP programs. In fact, there exists numerous examples of systems designed to measure performance across procedures, hospitals, and even clinicians themselves [7]. While such measures have traditionally been viewed as a tool to manage performance and quality improvement, the attachment of financial incentives to performance metrics has situated these tools as a cornerstone in healthcares value-based paradigm shift [7], [8].

As with many comparative measures, the adoption of these rankings has produced a strong desire by those being evaluated to discern specific factors influencing the underlying metric values. In doing so, researchers have predominantly focused on intrinsic elements of hospital care, including procedure complexity or characteristics of patients and their condition [9], [10]. Yet, the healthcare community was quick to point out this approach overlooked the socioeconomic and lifestyle factors known to significantly influence an individual's health and wellness, deemed their social determinants of health [11].

Collectively, these voices have echoed a British Medical Journal editorial, lamenting that "if deprived areas are not to be penalized for poor performance then the data must be adjusted to account for socioeconomic factors" [12]. As a result, recent works expanded their purview to include social elements from poverty levels to an individual's access to clinicians. In doing so, they have demonstrated ties between community factors and notable performance metrics such as readmission rates. In one study reporting that "Fifty-eight percent of national variation in hospital readmission rates was explained by the county in which the hospital was located" [13].

Yet, despite such a call to action, performance measurement frameworks continue to utilize measures that beget this type of penalization. The implications such oversight can be seen in relation to the currently discontinued CMS Star Rating metric. Designed to provide an assessment of a hospital quality, a landmark study published in JAMA found strong associations between stress levels in a city, and hospital rankings [14].

The work presented in this manuscript builds on this notion. Extending beyond overall quality, to explore novel relations between specific HVBP performance domains and health factors of the community. In doing so we also expand on the set of health factors utilized by prior works, demonstrating how a diverse array can be drawn from across multiple publicly available sources, In turn highlighting how such factors can offer more comprehensive analysis and deeper insights into external factors associated with hospital performance domains.

III. Data

A. CMS Data

At the foundation of this study are performance measures drawn from the 2018 HVBP data, representing a national sample of hospitals participating in the Inpatient Prospective Payment System, a comprehensive set of inclusion and exclusion criteria for the program can be found in [5]. For each hospital, scores were extracted across four primary domains: 1) Clinical Care, 2) Patient-and Caregiver-Centered Experience of Care, 3) Safety, and 4) Efficiency and Cost Reduction.

Of note, while reimbursements are computed through a complex formulation of achievement and improvement scores, this study focused on the underlying raw scores of each domain. These measures are preferable as they form the basis of achievement scores used to compare between hospitals; offing an objective value on which to analyze community relations. Additionally, raw measures help limit the introduction of bias as improvement scores are derived from historical performance; potentially confounding the any identified relations.

B. Community Data

Next, we set out to capture a detailed view of the community in which each hospital resides. Doing so required us to look beyond any single agency or data repository, ultimately bringing together three disparate datasets from government and academic sources. A brief summary of each is provided below, while a complete set of features can be found in Fig. 1.

First, from the American Community Survey provided by the U.S. Census Bureau we obtained population estimates for a range of demographic and occupational attributes. Moving along, we utilized the Food Environment Atlas, developed by U.S. department of agriculture (USDA) [15]. Although often overlooked as health data source, the atlas provided granular data around food access and social program utilization. Lastly, we turned to an academic dataset from the University of Wisconsin known as the County Health rankings (CHR). The CHR was designed to capture an expansive set of factors around the "social, economic, physical, clinical, and other factors that influence both how long and how well we live" [16].

C. Data Linking and Cohort Selection

Prior to performing any analysis, we selected a subset of hospitals to minimize known sources of bias. First, with respect to the HVBP data, hospitals that did not report scores for all performance domains were removed. As we did not wish to distort the analysis of community factors with latent associations to a hospitals compliance with reporting standards. Similarly, those hospitals with an efficiency score of 0 were also removed. As the efficiency domain score represents an aggregate of only a single sub-score, we did not want to penalize hospitals that could not or did not report that metric.

Ultimately, the community data was linked to the HVBP performance scores utilizing data crosswalks matching hospital zip codes provided by CMS with the ZCTA, and FIPS values employed by the USDA, CHR, and Census Bureau. In total, our data captured performance and socioeconomic information for an expansive set of 1,426 hospitals across the US.

IV. Methods

The analyses in our study can broadly be broken into two parts. First, an establishment of scoring profiles derived from the HVBP performance measures, and second a model-based analysis of community factors differing between such profiles.

Census Bureau: American Community Survey (ACS)	University of Wisconsin & RWJF: County Health Rankings (CHR)	USDA: Food Environment Atlas	CMS: Hospital Value-Based Purchasing
• Total population (HC01_EST_VC01) • Median age years (HC01_EST_VC12) • % Sex: Male (HC01_EST_VC15) • % Civilian pop. with health insurance (HC03_VC131) • Civilian Unemployment Rate (HC03_VC12) • Median income (HC01_EST_VC63)	• Poor or fair health (chr_2) • Primary care physicians (chr_4) • Preventable hospital stays (chr_5) • Diabetes prevalence (chr_60) • Air pollution particulate matter (chr_125) • Severe housing problems (chr_136) • Drug overdose deaths (chr_138)	• Fast-food restaurants/1,000 pop (FFRPTH) • % pop. Household food insecurity (FOODINSEC) • % Households, no car & low access to store (LACCESS_HHNV) • Recreation & fitness facilities/1,000 pop (RECFACPTH) • % pop. National School Lunch Program participants (NSLP) • % pop. WIC participants (WIC)	• Clinical Care • Safety • Efficiency and Cost Reduction • Patient - and Caregiver-Centered Experience of Care / Care Coordination

Fig. 1. Data elements and (abbreviation codes) utilized for analysis, broken down by source and agency

A. Hospital Performance Groupings

To begin, it is important to remember that the HVBP program is comprised of four scoring domains, each capturing various aspects of hospital care and operations. However, as all metrics arise from the same hospital, it is likely latent relations exist between high and low scores in each domain. As such, it is inadequate to simply isolate similarly performing hospitals across any single domain. Rather, a more rigorous evaluation necessitated the identification of similar hospitals across the four domains; identifying specific combinations of scores that create well defined subgroups within the broader set of hospitals. To do so, we employed K-means clustering, utilizing the gap statistic to identify an optimal K of 4.

B. Association of Community Factors and Scoring Profiles

Utilizing the four scoring profiles, we next moved to identify statistical relations between such groupings and the community factors of the hospitals within them. To do so we employed a One Vs. Rest (OvR) approach, identifying factors that differentiate one profile from remaining three.

As it is likely a scoring profile would be reflective of a combination of community factors, we employ L1 regularized logistic regression for all analyses. Offering a systemic method to adjust for multiple features while producing a measure of effect size through the coefficients.

Further, given the interconnected nature of community factors, consideration was given during feature selection to avoid multicollinearity, and thus provide more reliable inference around coefficient significance. Bivariate correlations were computed between all factors using the non-parametric Spearman's statistic, and pairs with a correlation above .65 were removed. Note: Fig. 1 represents the final feature set.

V. Results and Discussion

A. Hospital Performance Groupings

Looking first to scoring profiles 1-4, we find the average total performance the hospitals to be uninformative, with values of 71.2 (n:276), 65.0 (n:256), 21.4 (n:466), and 53.4 (n:428) respectively. However, the mean performance across the four HVBP scoring domains of each profile tells a different story. Seen in Fig. 2, it is highly encouraging to observe well-defined subpockets comprised of high and low *combinations* across the domains; supporting the need to cluster such values to capture relations among the scores.

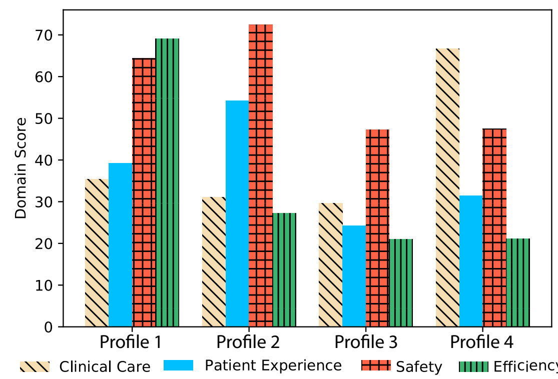

Fig. 2. Mean domain performance scores by profile

To quantifiably demonstrate these groupings added value beyond simply partitioning high and low performing hospitals we took the analysis a step further, computing the Jaccard similarity between the set of hospitals in each profile and the within each quantile of each HVBP scoring domain independently. Among the 16 possible sets, the highest overlap for any pair was 39%; illustrating clear value in accounting for combinations of scoring domains for hospital groupings.

B. Association of Community Factors and Scoring Profiles

We next turned to the analysis of how hospital community factor differ amongst the scoring profiles. Fig. 3 presents a comprehensive overview of the OvR model analysis, where those cells in grey represent factors with no significant difference between the designed profile and the remaining three. While those with shading designate significant differences at $p < .05$. To further aid in interpretation, the respective label of each such cell offers the adjusted odds ratio.

Looking to set of significant factors with respect to each profile, we find several intriguing relations emerge between the socioeconomic and health conditions of a hospitals community, and the HVBP scoring domains comprising the profile. Due to space limitations, we focus this discussion on two specific profiles, in which we support veracity of the identified relations with existing healthcare literature. Note, for interpretability, we concentrate on factors with odds ratios >1. As

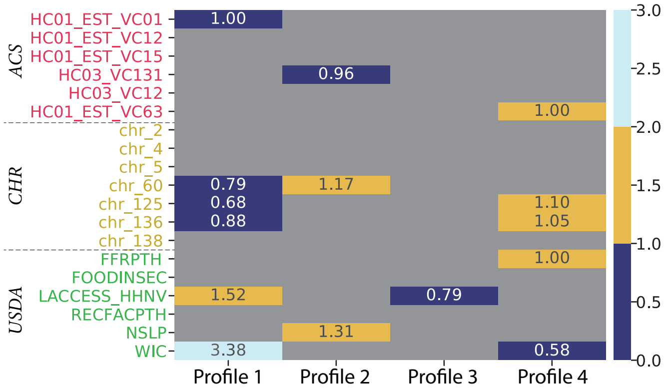

Fig. 3. Association of community factors to scoring profiles. Results are read column-wise, identifying factors differentiating the respective profile from the remaining three. Key — *Grey*: No significant association, *Shaded*: Significant at p<.05, *Label*: Adjusted Odds Ratio

these factors demonstrate an increased likelihood of occurring in the respective scoring profile, while those <1 represent likelihood within the aggregate of the remaining three.

1) Profile 2: Characterized by lower clinical and efficiency scores, we find hospitals exhibiting profile 2 tend to occur in communities with increased prevalence of diabetes, and children in the National School Lunch Program (NSLP).

As we explore these factors in the context of the HVBP, it is important to remember that each scoring domain is comprised of several sub-scores. In the case of Clinical Care domain, these include mortality rates of acute myocardial infarction and heart failure. Both of which have extensive evidence tying increased mortality rates to patient's diabetic status [17].

In a similar fashion, we again find support relating participation in NSLP to lower average efficiency scores in the sub-scores of the domain. Eligibility for the NSLP is determined in part by family income relative to the poverty level, and for 2018, the efficiency score was computed using a single sub-score, the "Medicare spending per beneficiary." A factor previous works have established associations with the poverty levels of both individuals and communities [18].

2) Profile 4: Moving next to the scoring profile associated with profile 4, our results indicate an increased likelihood of hospital communities exhibiting higher air pollution rates, and severe housing issues. Additionally, although lower in effect size, there exist significant relations in average median income and fast food resultants per 1000/pop. Together, such community characteristics align well with those of metropolitan areas.

It is understandable then to find such factors associated with the highest average clinical and safety performance scores amongst the scoring profiles. As Lutfiyya et al. demonstrated numerous differences in clinical performance between urban and rural hospitals [19]. Moreover, we note profile 4 also exhibits lower patient experience scores. A finding in line with a recent study illustrating that significant variance in patient experience scores can be explained by accounting for the region of the country where care was provided [20].

VI. Conclusion

Viewed in their entirely, the relations identified by this study offer strong evidence supporting the hypothesis that performance domains utilized by the HVBP program do not exist in isolation, but are intertwined with the socioeconomic and health factors of the communities in which a hospital resides. However, acknowledging such relationships is only the first step. As ranking systems become increasingly entrenched within the U.S. health system, there exists a clear need to broaden the factors by which quality measures are computed.

Much as we have learned to risk-adjust hospitalizations, work remains to understand how the influences of community factors can be accounted for within performance metrics. It is our hope this work serves as a foundation to pursue such extensions. Demonstrating how public use data provide information into factors outside of the hospital walls, and deeper insights into specific dimensions of performance scores. Together this information holds promise to address research questions aimed to advance and improve our health system.

References

[1] D. M. Berwick, T. W. Nolan, and J. Whittington, "The triple aim: care, health, and cost," *Health affairs*, vol. 27, no. 3, pp. 759–769, 2008.

[2] H. Fröhlich *et al.*, "From hype to reality: data science enabling personalized medicine," *BMC medicine*, vol. 16, no. 1, p. 150, 2018.

[3] H. D. Miller, "From volume to value: better ways to pay for health care," *Health Affairs*, vol. 28, no. 5, pp. 1418–1428, 2009.

[4] T. T. Chee *et al.*, "Current state of value-based purchasing programs," *Circulation*, vol. 133, no. 22, pp. 2197–2205, 2016.

[5] CMS, "Medicare program; hospital inpatient value-based purchasing program. final rule." *Federal register*, vol. 76, no. 88, p. 26490, 2011.

[6] J. M. VanLare *et al.*, "Value-based purchasing-national programs to move from volume to value," *NEW ENGL J MED*, vol. 367, 2012.

[7] V. A. Kazandjian and T. Lied, *Healthcare performance measurement: systems design and evaluation*. ASQ Quality Press, 1999.

[8] O. A. Arah *et al.*, "Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement," *International journal for quality in health care*, 2003.

[9] S. K. Pasquali *et al.*, "Impact of patient characteristics on hospital-level outcomes assessment in congenital heart surgery," *The Annals of thoracic surgery*, vol. 100, no. 3, pp. 1071–1077, 2015.

[10] J. L. Paruch *et al.*, "Impact of hepatectomy surgical complexity on outcomes and hospital quality rankings," *Annals of surgical oncology*, vol. 21, no. 6, pp. 1773–1780, 2014.

[11] M. Marmot, "Social determinants of health inequalities," *The lancet*, vol. 365, no. 9464, pp. 1099–1104, 2005.

[12] J. Mulligan *et al.*, "Measuring the performance of health systems: indicators still fail to take socioeconomic factors into account," 2000.

[13] J. Herrin *et al.*, "Community factors and hospital readmission rates," *Health services research*, vol. 50, no. 1, pp. 20–39, 2015.

[14] J. Hu and D. Nerenz, "Relationship between stress rankings and the overall hospital star ratings: an analysis of 150 cities in the united states," *JAMA internal medicine*, vol. 177, no. 1, pp. 136–137, 2017.

[15] Economic Research Service (ERS), U.S. Department of Agriculture (USDA), "Food environment atlas."

[16] P. L. Remington *et al.*, "The county health rankings: rationale and methods," *Population health metrics*, vol. 13, no. 1, p. 11, 2015.

[17] R. Dhingra and R. S. Vasan, "Diabetes and the risk of heart failure," *Heart failure clinics*, vol. 8, no. 1, pp. 125–133, 2012.

[18] J. Cubanski *et al.*, "The latest on geographic variation in medicare spending: a demographic divide persists but variation has narrowed," *Kaiser Family Foundation*, 2015.

[19] M. N. Lutfiyya *et al.*, "A comparison of quality of care indicators in urban acute care hospitals and rural critical access hospitals in the united states," *International journal for quality in health care*, vol. 19, 2007.

[20] J. Herrin *et al.*, "HCAHPS scores and community factors," *American Journal of Medical Quality*, p. 1062860618765977, 2018.