Variation in Batch Ordering of Imaging Tests in the Emergency Department and the Impact on Care Delivery

Jacob Jameson, MS[1,2]

Soroush Saghafian, PhD, MS[2]

Robert S. Huckman, PhD[3,5]

Nicole Hodgson, MD[4]

[1]Interfaculty Initiative in Health Policy, Harvard University, Cambridge, MA, USA
[2]Harvard Kennedy School, Harvard University, Cambridge, MA, USA
[3]Harvard Business School, Harvard University, Boston, MA, USA
[4]Mayo Clinic, Department of Emergency Medicine, Phoenix, AZ, USA
[5]National Bureau of Economic Research, Cambridge, MA, USA

Code Repository: https://github.com/jacobjameson/batch-vs-sequence/

Abstract

Importance: Prior research has demonstrated that test ordering can significantly impact Emergency Department throughput, cost, and quality of care, yet there is no consensus on the optimal test ordering strategy.

Objectives: To examine heterogeneity in physician batch ordering practices and measure the associations between a physician's tendency to batch order imaging tests on patient outcomes and resource utilization.

Design, Setting, and Participants: In this retrospective study, we used comprehensive EMR data from patients who visited the Mayo Clinic of Arizona ED between 10/6/2018 and 12/31/2019. The ED recorded approximately 50,836 visits, all randomly assigned to physicians during the study period. Patients were managed across 26 treatment rooms and up to 9 hallway spaces.

Main Outcomes and Measure: Primary outcomes are patient length of stay (LOS) in the ED, number of diagnostic imaging tests ordered during a patient encounter, and patients' rate of return to the ED within 72 hours. The association between outcomes and physician batch tendency was measured using a multivariable linear regression controlling for various covariates.

Results: Our analysis reveals a significant positive association between a physician's tendency to batch order imaging tests and an increased LOS. Specifically, having a physician with a batch tendency 1SD greater than the average physician was associated with a 4.5% increase in ED LOS ($p < 0.001$). It was also associated with an 8.29% (0.3 percentage points) decrease in the probability of a 72-hour return ($p < 0.001$), implying that batching may lead to more comprehensive evaluations, reducing the need for short-term revisits. A batch tendency 1SD greater than that of the average physician was also associated with an additional 8 imaging tests ordered per 100 patient encounters ($p < 0.001$), suggesting that batch ordering may be leading to tests that would not have been otherwise ordered had the physician waited for the results from one test before placing their next order.

Conclusions and Relevance: This study highlights the considerable impact of physicians' diagnostic test ordering strategies on ED efficiency and patient care. Our findings indicate that, on average, sequential ordering of tests enables physicians to serve patients more efficiently by using the information obtained from prior tests (an information gain advantage). The results also highlight the need to develop guidelines to optimize ED test ordering practices.

Introduction

Emergency departments (EDs) serve as critical junctures in healthcare delivery, balancing the immediate needs of patients with the overarching operational and administrative demands of hospital management. This balance is precarious and affected by numerous factors, including the strategic ordering of diagnostic imaging tests—a common yet complex practice with implications for patient flow, hospital costs, and patient safety[1]. The efficiency of the ED is not just a matter of patient care but also a significant hospital management concern, with the potential to influence hospital-wide operational dynamics and financial health[2,3].

An understudied aspect of ED efficiency is the practice of batch ordering imaging tests. Given the long turnaround times of imaging tests, a physician can ensure that their patient is in simultaneous waiting queues for each test by placing multiple orders simultaneously. While ostensibly a measure to expedite patient diagnosis and treatment, batch ordering raises several potential concerns. For instance, the case of a patient presenting with nonspecific abdominal pain could lead to a batch order, including an abdominal CT scan, ultrasound, and X-ray. While comprehensive, this approach raises questions about the necessity of each test, the patient's cumulative radiation exposure, and the impact on the patient's length of stay and overall healthcare costs[4-8].

Furthermore, the financial implications extend beyond the cost of the tests themselves. Though sometimes necessary for thorough evaluation, an increased length of stay can also contribute to hospital overcrowding and reduced capacity for new patients, exacerbating operational pressures and financial constraints on the healthcare system[9]. This delicate balance between ensuring rapid, accurate diagnosis and minimizing unnecessary use of resources is a central challenge in hospital management, reflecting broader concerns about the sustainability of healthcare practices[10].

Despite its significance, the impact of batch ordering on these dimensions remains underexplored. The assumption that batch ordering represents an efficient test ordering practice has not been rigorously examined, leaving a gap in our understanding of its true operational and economic implications. This study aims to shed light on this critical issue, exploring how batch ordering of imaging tests affects the length of stay, total testing volume (surrogates for efficiency), and the need for short-term revisits (a surrogate for effectiveness).

By situating this investigation within the context of hospital management, we seek to determine whether the perceived efficiency of batch ordering aligns with its actual outcomes, providing evidence-based insights that can guide future policy and practice in emergency care[11,12].

Methods

Study Design and Setting: Our retrospective observational study was conducted in the Mayo Clinic of Arizona ED. During the study period, the ED recorded 50,836 visits, managed across 26 treatment rooms and up to 9 hallway spaces. The department is exclusively staffed by board-eligible or board-certified emergency physicians (EPs), with rotating residents overseeing about 10% of patient volume. Physicians operate in a unique workflow that includes staggered 8.5-hour shifts and a randomized assignment system that reduces systematic differences in patient populations served by different physicians[13].

We retrospectively reviewed comprehensive ED operational data from 10/6/2018 through 12/31/2019. The dataset includes detailed patient demographics, chief complaints, vital signs, emergency severity index (ESI), length of stay (LOS), and resource utilization metrics. This period was chosen to provide a robust data set while excluding the influence of the coronavirus pandemic. We further restricted our sample to patient encounters serviced by full-time physicians and broad chief complaint areas seen in over 1,000 encounters over the study period (i.e., excluding rare complaints). The final sample included 43,299 patient encounters and contained no missing data for covariates used in the analysis.

Details on Data: A critical aspect of our data is the random patient-to-physician assignment. In most EDs, physicians have some discretion in selecting the patients they see from the pool of those waiting for treatment. In contrast, patients arriving at the Mayo Clinic ED are assigned to physicians via a randomized rotational patient assignment algorithm, which practically removes potential selection bias concerns from our analyses[12]. In essence, barring arrival time and shift-level variation, the physician-to-patient matching can be deemed random. Table 1 in the Results section confirms that the complaints and severity of patients served are balanced across physicians.

Measurements: We define "batching" in line with standard emergency medicine practices. Batching occurs when a physician simultaneously orders a comprehensive set of diagnostic tests, typically covering a broad range of potential diagnoses. This contrasts with sequential ordering, where tests are ordered sequentially based on the information obtained from each test as needed.

We operationalize batching as occurring when multiple diagnostic imaging tests are ordered within a 5-minute window. Sensitivity analyses around this cutoff point (eTable 1) show that our results are robust to this definition. Each imaging test (e.g., X-ray, CT scan) is considered a separate, distinct test for our study. Therefore, a batch in our study consists of two or more distinct imaging tests.

Statistical Analysis: To assess the impact of batching on various outcomes of interest, we developed a measure to quantify each physician's tendency to batch. This "batch tendency" score allows us to explore the associations between batching behavior and critical outcomes such as patient length of stay, resource utilization, and 72-hour return to the ED. The batch tendency for each physician was calculated using a "leave-one-out" approach. Specifically, we estimate the following multivariable logistic regression for each patient encounter:

$$logit\left(P_{Batched_{i,t}}\right) = \beta_0 + \beta_1 \boldsymbol{X}_{ym} + \beta_2 \boldsymbol{X}_{dt} + \beta_3 \boldsymbol{X}_{complaint/severity} + \beta_4 \boldsymbol{X}_{physician\ ID} + \varepsilon_{i,t}$$

Where $Batched_{i,t}$ is a dummy variable equal to one if patient i had their imaging tests batch ordered on the encounter that took place on date t. Covariates include year-month, $\boldsymbol{X}_{ym}$, to control for time and seasonal variation in batching, such as hospital-specific policies (e.g., initiatives to eliminate excess testing) or seasonality in ED visits. We also control for shift-level variations that include physician scheduling and patient arrival with day-of-week and time-of-day covariates, $\boldsymbol{X}_{dt}$. Chief complaint by severity, $\boldsymbol{X}_{complaint/severity}$, is also included to increase precision. As stated earlier, these controls are required for our quasi-random assignment assumption. We use this model to produce a predicted probability of batching occurring for each patient encounter.

For physician j serving patient i, we then compute the leave-one-out average of $P_{Batched_{i,t}}$ for each physician j by excluding the current patient i from the calculation and including all other patients served by physician j during the study period. This measure effectively eliminates the mechanical bias resulting from patient $i's$ own case influencing the physician's batch tendency score and captures the physician's general likelihood of batching imaging tests across a wide range of cases[14,15].

After calculating each physician's average leave-one-out batch tendency, we center and standardize it into a uniform scale, facilitating more straightforward interpretation and comparison across physicians. Figure 1 shows the relationship between batch tendency and batch ordering for each specific patient encounter. This strong relationship between batch tendency and batch ordering allows us to think of batch tendency as an instrumental variable (IV), which addresses the problem of endogeneity[1] in studying the impact of batching[16]. This enables us to use batch tendency as a proxy for batching itself. We expect that a physician's batch tendency affects our primary outcomes only through its influence on actual batching behavior. In other words, a physician's tendency to batch should not directly impact patient length of stay, 72-hour return rate, or the number of imaging tests ordered, except through the actual practice of batch ordering.

All statistical analyses were performed using R (version 4.3.2). All multivariable linear regression models control for calendar month and time-of-day fixed effects, which is necessary to achieve quasi-random assignment. We additionally control for patient chief complaint and severity, an indicator for whether laboratory tests were ordered for the patient, and hospital occupancy level to improve precision. We use robust standard errors clustered at the physician level.

We evaluate the influence of physicians' batch ordering tendency on three patient-level dependent variables: length of stay (LOS), the 72-hour return rate, and the number of distinct imaging tests ordered. Additional analyses examine interactions between batch tendency and other key variables, such as patient complaint and ESI, to explore whether the effect of batch ordering varies across different patient acuities and complaints. Because our data regarding 72-hour returns are limited to returns to the same ED, we expect that the magnitude of our estimate is biased towards the null.

[1] Endogeneity occurs when an explanatory variable is correlated with the error term, leading to biased estimates. It can arise due to omitted variables, measurement error, or simultaneity.

Figure 1: Relevance of Batch Tendency on Batch Ordering Probability

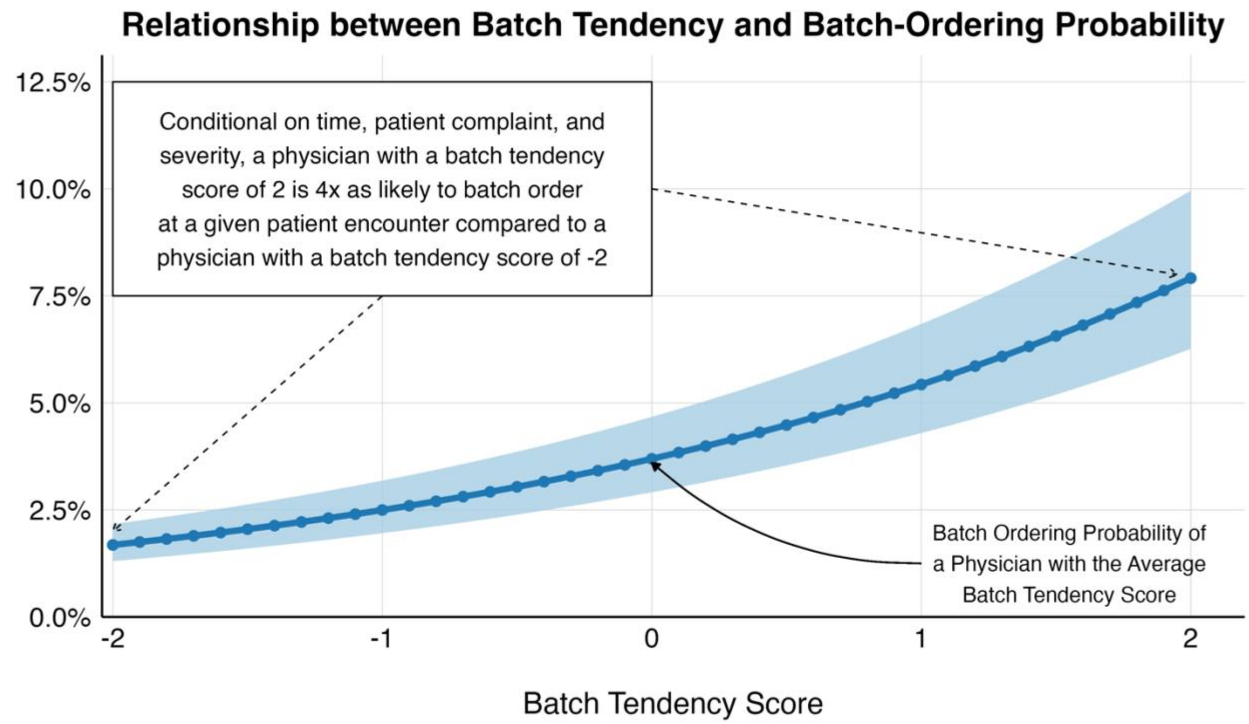


Figure 1 shows the predicted probability of batch-ordering at a given patient encounter, conditional on time, patient complaint, and severity, from a logistic regression model controlling for these features. The x-axis represents the batch tendency score, which measures the physician's tendency to batch-order. The red line represents the predicted probability of batch-ordering at a specific patient encounter, and the shaded area represents the 95% confidence interval.

Data Manipulation: As evidenced in the literature, transforming the outcome variable can improve the performance of regression models. For right-skewed outcomes, such as the length of stay (LOS), which is shown to be log-normal, applying a natural log transformation can lead to a more symmetric distribution and mitigate the influence of outliers[17–19]. This approach is analogous to count data models, such as Poisson or negative binomial regression, which are designed to handle skewed distributions inherent in count data and have been used previously[18]. As demonstrated in eFigure 1, the distribution of LOS in our data is highly right-skewed. We thus apply a natural log transformation to this variable before it is used in our regression analyses. We report the un-exponentiated coefficients from these models in Table 2, which can be interpreted as a $100 \times (e^{\beta} - 1)$ percent change in LOS for a given 1 unit increase in our independent variable of interest, where β is the coefficient on our independent variable of interest.

Results

Table 1 displays the results of a Wald balance test, showing that complaints and severity of patient encounters are balanced across physicians in our study. In other words, due to the random assignment, all physicians served a similar portfolio of patients regarding presenting complaints and severity. This is a critical aspect of our study, as it ensures that differences in test ordering behavior are attributable to physician practice patterns rather than patient characteristics.

Table 1: Balance Test for Random Assignment

Chief Complaints	**Frequency No. (%)**	**F-Statistic**	***p-value***
Abdominal Complaints	6,232 (14%)	1.401	0.095
Back or Flank Pain	2,550 (6%)	1.029	0.423
Chest Pain	3,521 (8%)	1.042	0.406
Extremity Complaints	5,259 (12%)	0.991	0.472
Assaults and Trauma	2,381 (5%)	0.773	0.769
Gastrointestinal Issues	3,323 (8%)	1.027	0.425
Neurological Issue	3,492 (8%)	0.707	0.843
Shortness of Breath	2,962 (7%)	1.198	0.232
Skin Complaints	2,176 (5%)	1.021	0.433
Upper Respiratory Symptoms	1,915 (4%)	1.239	0.197
Emergency Severity Index (ESI)	**Frequency No. (%)**	**F-Statistic**	***p-value***
ESI 1 or 2	13,913 (32%)	1.277	0.169
ESI 3, 4, or 5	29,386 (68%)	1.277	0.169

Table 1 reports the results of a Wald test, which was conducted to assess the balance of chief complaints across physicians in our dataset. We created chief complaint categories before analysis by grouping similar presenting issues. Vital signs were categorized as follows: tachycardia (pulse more significant than 100), tachypnea (respiratory rate greater than 20), fever (temperature greater than 38∘C), and hypotension (systolic blood pressure less than 90). A balanced distribution implies that complaints and severity are evenly distributed across physicians, which we expect to be the case due to randomization. The Wald F-statistic and p-value are reported. Robust standard errors (type HC1) accounted for potential heteroscedasticity in the data.

The data also indicate differences in physician batch ordering rates across complaint categories (Figure 2). Notably, the variation in batching was most pronounced during patient encounters where the presenting complaint was neurological or trauma related. We note that at least one imaging test was ordered in approximately 31,498 of the 43,299 patient encounters in our study. While only 2,421 (7.7%) of these encounters involved image batching, 7,181 (22.8%) of the non-batched encounters resulted in the physician ordering at least one more imaging test after placing the first order.

Figure 2: Variation in Physician Imaging Batch Rates

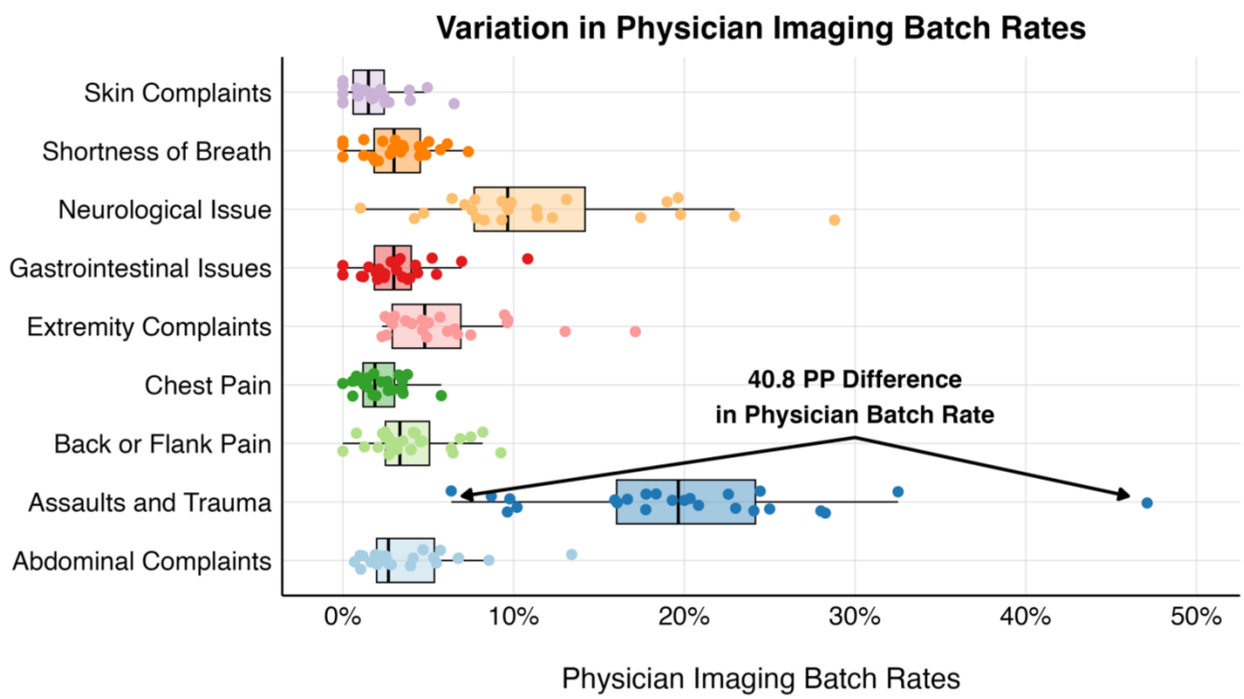


Figure 2 illuminates the marked differences among physicians in their propensity to batch order imaging tests. The 24 physicians are represented with points, revealing that specific complaint areas have more variance than others regarding differing batch rates among physicians.

Table 2 presents the linear regression coefficients for the impact of batch tendency on three primary outcomes: the natural logarithm of ED length of stay (ln(LOS)), 72-hour return, and the number of distinct imaging tests ordered. Our analysis reveals a significant positive association between a physician's tendency to batch order imaging tests and an increased ln(LOS), with a coefficient of 0.045 (95% CI = [0.005, 0.084], $p < 0.001$). This implies that having a physician with a batch tendency 1SD greater than the average physician is associated with a 4.5% increase in ED length of stay. However, we also find that a batch tendency 1SD greater than the average physician is associated with an 8.29% (0.3 percentage points) decrease in the probability of a 72-hour return, indicated by a coefficient of -0.003 (95% CI = [-0.005, -0.001], $p < 0.001$), implying that batching may lead to more comprehensive initial evaluations, reducing the need for short-term revisits. Finally, there is a notable association with an increased number of distinct imaging tests ordered, as evidenced by a coefficient of 0.08 (95% CI = [0.066, 0.094], $p < 0.001$), underscoring that batch ordering may be leading to tests that would not have been otherwise ordered had the physician waited for the results from one test before ordering the next.

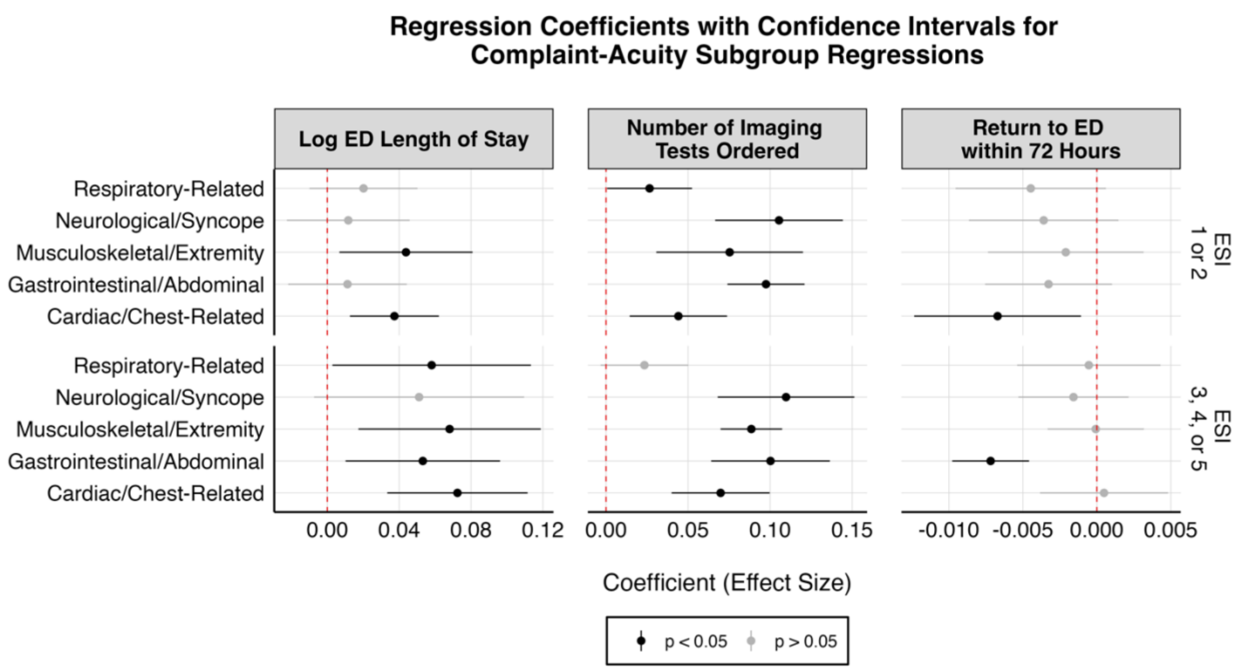
Table 2: Main Multivariable Regression Results of Primary Outcomes on Batch Tendency

	Dependent Variables		
	ln(LOS)	72 Hour Return	Number of Distinct Imaging Tests
Batch Tendency	0.045^{***}	-0.003^{***}	0.080^{***}
	(0.020)	(0.001)	(0.007)
Controlling for time and shift?	Yes	Yes	Yes
Controlling for complaints and ESI?	Yes	Yes	Yes
Controlling for hospital occupancy?	Yes	Yes	Yes
Controlling for lab tests ordered?	Yes	Yes	Yes
Observations	43,299	43,299	43,299

The coefficient comes from a multivariable linear regression where we regress batch tendency on our primary outcomes. We control for time and shift fixed effects (necessary for quasi-random assignment), patient-level variables, hospital occupancy, and whether the patient also had laboratory tests ordered during their visit. Standard errors are clustered at the physician level.
**p<0.1; **p<0.05; ***p<0.01*

Figure 3 displays the results of the subgroup analysis stratified by the patient's Emergency Severity Index (ESI) and broad patient complaint category (as defined in eTable 2 and eTable 3). Results indicate heterogeneity in the effect of batch tendency across patient complaints and acuity. Notably, among the most acute patients (ESI 1 and 2), the propensity to batch order image tests was generally associated with significant increases in LOS and the total number of imaging tests ordered. We do not see substantial reductions in the probability of a 72-hour return (though this coefficient is biased towards the null), except for the case of high acuity patients presenting with cardiac/chest-related complaints and lower acuity patients presenting with gastrointestinal/abdominal complaints. For these subgroups, a higher batch tendency, which is associated with a greater testing volume, appears to decrease the likelihood of short-term readmission. This suggests that for these types of patients, the benefits of a more comprehensive initial evaluation through batch ordering may outweigh the potential drawbacks observed in other subgroups[20,21].

Figure 3: Regression Coefficients with Confidence Intervals from Subgroup Analysis



The coefficient comes from a multivariable linear regression where we regress batch tendency on our primary outcomes for each complaint by acuity subgroup. We control for time and shift fixed effects (necessary for quasi-random assignment), patient-level variables, hospital occupancy, and whether the patient also had laboratory tests ordered during their visit. Standard errors are clustered at the physician level

Discussion

Our study highlights that patterns of diagnostic test ordering in the ED have profound implications on the efficiency of care delivery and patient outcomes. Our investigation of the variation in test ordering behaviors within a controlled ED environment brings to light the need for targeted diagnostic strategies over a one-size-fits-all approach. Our findings contribute to the growing body of evidence supporting the use of data-driven, personalized approaches in ED management. This aligns with the broader shift towards evidence-based medicine, which emphasizes the integration of best research evidence, clinical expertise, and patient values to optimize care delivery and system efficiency[22]. We can develop more effective, evidence-based strategies for ED resource utilization and patient management by understanding the nuances of test ordering practices and their impact on patient outcomes.

The physician-level variability in inclination towards batching and non-batching test orders—within the same ED environment—raises fundamental questions about the underpinnings of clinical decision-making. Notably, our study revealed that non-batchers, who potentially employ a more judicious and sequential approach to ordering tests, could achieve a shorter length of stay (LOS) without negatively impacting the 72-hour return rates for a large subset of patients. This is due to the information gain advantage of sequential test ordering, where the results of one test may eliminate the need for another. This result aligns with previous research emphasizing the importance of tailored diagnostic pathways in achieving optimal health outcomes and operational efficacy[23–26].

Over-testing in EDs is not a benign phenomenon. It is associated with increased risks, including patient exposure to unnecessary radiation and the resultant psychological and physical burden from incidental findings[27]. Moreover, the economic implications are substantial, with the overuse of diagnostic tests contributing significantly to the escalating costs of healthcare[28]. As such, our results suggest the need to examine the practice of batching across different clinical conditions and in other clinical settings beyond the ED[29].

Incorporating physician test ordering tendencies into ED management strategies is complex but potentially beneficial. Recent initiatives have experimented with optimizing patient-physician matching based on various factors, including patient complaints and physician expertise[30]. Our findings suggest that considering physicians' test ordering tendencies, alongside these other factors, could help strike a balance between ensuring thorough patient evaluation and minimizing unnecessary resource utilization. By aligning physician test ordering behaviors more closely with patient needs, EDs may enhance patient satisfaction and outcomes while improving operational efficiency[24].

Future studies should investigate the subtleties of the information gain advantage from sequential testing versus the potential benefits of batching. There is a delicate balance between thoroughness and efficiency, which becomes even more precarious in high-stakes environments such as the ED. Understanding and navigating this balance could yield significant advancements in patient care and ED operations.

Limitations

Our study involves multiple considerations that may limit the interpretation and application of our findings. While our data involve random assignment of patients to physicians, the variation we observe across physicians could stem from myriad sources, including physician training, accumulated experience, and general inclinations toward more testing[28]. These influences could drive a physician toward a particular testing methodology, confounding the batch tendency measure with other characteristics of the physician's approach to practice. Furthermore, though we consider ED physicians to be independent actors, it is known that they affect each other's speed and quality[31]. Therefore, moving beyond associative insights is imperative as research in this area of inquiry progresses.

We also acknowledge the possibility of Type I error due to multiple comparisons in our subgroup analysis. In this analysis, we conducted multiple hypothesis tests across different patient complaints and acuity subgroups. As the number of hypothesis tests increases, the likelihood of observing a statistically significant result by chance increases, potentially leading to false positives. Given this limitation, our subgroup analyses should be considered exploratory and interpreted cautiously.

Finally, the generalizability of our results may be limited due to the study's single-site design. The Mayo Clinic's operational procedures, patient demographics, and physician culture may not reflect those of other EDs, potentially affecting external validity.

Conclusion

Our study contributes to a critical conversation on optimizing diagnostic processes in the ED. It underscores the need for diagnostic strategies that consider the heterogeneity in the effects of batch ordering across different clinical conditions to enhance operational efficiency and quality of care. Through a detailed investigation of batch ordering practices and their comparison to non-batch ordering, which includes both sequential testing and single-test or no-test strategies, we have highlighted the implications of these behaviors on patient length of stay, resource utilization, and hospital costs, providing new evidence that challenges the current paradigm of diagnostic testing in the ED.

Notably, batching was associated with increased ED lengths of stay and testing, with a small but significant average decrease in the probability of a 72-hour return. Beyond these average results, our heterogeneity analyses suggested that a more discerning approach to test orders can be efficient and beneficial to patient outcomes. These findings echo the call for targeted approaches to be integrated into ED operations, emphasizing the importance of tailoring diagnostic processes to individual patient needs to avoid over-testing pitfalls, such as unnecessary radiation exposure and the financial burden of healthcare delivery.

References

1. Cournane S, Conway R, Creagh D, Byrne D, Sheehy N, Silke B. Radiology imaging delays as independent predictors of length of hospital stay for emergency medical admissions. *Clinical Radiology*. 2016;71(9):912-918. doi:10.1016/j.crad.2016.03.023

2. Darraj A, Hudays A, Hazazi A, Hobani A, Alghamdi A. The Association between Emergency Department Overcrowding and Delay in Treatment: A Systematic Review. *Healthcare*. 2023;11(3):385. doi:10.3390/healthcare11030385

3. Hodgson N, Saghafian S, Klanderman M, Urumov A, Traub S. Physician-driven early evaluation: Encounters seen in a vertical model. *JEM Reports*. 2021;2(2):100028.

4. Tamburrano A, Vallone D, Carrozza C, others. Evaluation and cost estimation of laboratory test overuse in 43 commonly ordered parameters through a Computerized Clinical Decision Support System (CCDSS) in a large university hospital. *PLoS One*. 2020;15(8):e0237159. doi:10.1371/journal.pone.0237159

5. Perotte R, Lewin G, Tambe U, others. Improving Emergency Department Flow: Reducing Turnaround Time for Emergent CT Scans. *AMIA Annual Symposium Proceedings*. 2018;2018:897-906.

6. Lyu H, Xu T, Brotman D, Mayer-Blackwell B, Cooper M, others. Overtreatment in the United States. *PLOS ONE*. 2017;12(9):e0181970. doi:10.1371/journal.pone.0181970

7. Jain S. Radiation in medical practice & health effects of radiation: Rationale, risks, and rewards. *Journal of Family Medicine and Primary Care*. 2021;10(4):1520-1524. doi:10.4103/jfmpc.jfmpc_2292_20

8. Traub S, Saghafian S, Judson K, et al. Interphysician Differences in Emergency Department Length of Stay. *Journal of Emergency Medicine*. 2018;54(5):702-710.

9. Sartini M, Carbone A, Demartini A, others. Overcrowding in Emergency Department: Causes, Consequences, and Solutions-A Narrative Review. *Healthcare*. 2022;10(9):1625. doi:10.3390/healthcare10091625

10. Ibanez M, Clark J, Huckman R, Staats B. Discretionary Task Ordering: Queue Management in Radiological Services. *Management Science*. 2017;64(9):4389-4407.

11. Saghafian S, Hopp W, Van Oyen M, Desmond J, Kronick S. Patient Streaming as a Mechanism for Improving Responsiveness in Emergency Departments. *Operations Research*. 2012;60(5):1080-1097.

12. Saghafian S, Austin G, Traub S. Operations Research/Management Contributions to Emergency Department Patient Flow Optimization: Review and Research Prospects. *IIE Transactions on Healthcare Systems Engineering*. 2015;5(2):101-123.

13. Traub S, Saghafian S, Judson K, et al. The Durability of Operational Improvements with Rotational Patient Assignment. *American Journal of Emergency Medicine*. 2018;36(8):1367-1371.

14. Eichmeyer S, Zhang J. Pathways into Opioid Dependence: Evidence from Practice Variation in Emergency Departments. *American Economic Journal: Applied Economics*. 2022;14(4):271-300.

15. Dobbie W, Goldin J, Yang C. The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges. *American Economic Review*. 2018;108(2):201-240.

16. Konetzka R, Yang F, Werner R. Use of instrumental variables for endogenous treatment at the provider level. *Health Economics*. 2019;28(5):710-716. doi:10.1002/hec.3861

17. Diehr P, Yanez D, Ash A, Hornbrook M, Lin D. Methods for analyzing health care utilization and costs. *Annual Review of Public Health*. 1999;20(1):125-144. doi:10.1146/annurev.publhealth.20.1.125

18. Cots F, Elvira D, Castells X, Sáez M. Relevance of outlier cases in case mix systems and evaluation of trimming methods. *Health Care Management Science*. 2003;6(1):27-35. doi:10.1023/A:1021908220013

19. Saghafian S, Kilinc D, Traub S. Dynamic Assignment of Patients to Primary and Secondary Inpatient Units: Is Patience a Virtue? *Working Paper, Harvard University (available at https://scholar.harvard.edu/saghafian/publications)*. Published online 2019.

20. Hodgson N, Saghafian S, Mi L, et al. Are Testers Also Admitters? Comparing Emergency Physician Resource Utilization and Admitting Practices. *American Journal of Emergency Medicine*. 2018;36(10):1865-1869.

21. Carpenter C, Raja A, Brown M. Overtesting and the Downstream Consequences of Overtreatment: Implications of "Preventing Overdiagnosis" for Emergency Medicine. *Academic Emergency Medicine*. 2015;22(12):1484-1492. doi:10.1111/acem.12820

22. Masic I, Miokovic M, Muhamedagic B. Evidence based medicine-new approaches and challenges. *Acta Informatica Medica*. 2008;16(4):219. doi:10.5455/aim.2008.16.219-225

23. Singh H, Graber M. Improving diagnosis in health care—the next imperative for patient safety. *New England Journal of Medicine*. 2015;373(26):2493-2495. doi:10.1056/NEJMp1512241

24. Saghafian S, Hopp W, Van Oyen M, Desmond J, Kronick S. Complexity-Augmented Triage: A Tool for Improving Patient Safety and Operational Efficiency. *Manufacturing and Service Operations Management*. 2014;16(3):329-345.

25. Lam J, Pickles K, Stanaway F, Bell K. Why clinicians overtest: development of a thematic framework. *BMC Health Services Research*. 2020;20(1):1011. doi:10.1186/s12913-020-05844-9

26. Saghafian S, Hopp W, Iravani S, Cheng Y, Diermeier D. Workload Management in Telemedical Physician Triage and Other Knowledge-Based Service Systems. *Management Science*. 2018;64(11):5180-5197.

27. Müskens J, Kool R, van Dulmen S, Westert G. Overuse of diagnostic testing in healthcare: a systematic review. *BMJ Quality & Safety*. 2022;31(1):54-63. doi:10.1136/bmjqs-2020-012576

28. Atkinson M, Saghafian S. Who should see the patient? On deviations from preferred patient-provider assignments in hospitals. *Health Care Management Science*. 2023;26:165-199.

29. Saghafian S, Hopp W. Can public reporting cure healthcare? The role of quality transparency in improving patient-provider alignment. *Operations Research*. 2020;68(1):71-92.

30. Abaluck J, Agha L, Kabrhel C, Raja A, Venkatesh A. The determinants of productivity in medical testing: Intensity and allocation of care. *American Economic Review*. 2016;106(12):3730-3764.

31. Saghafian S, Imanirad R, Traub S. Do Physicians Influence Each Other's Performance? Evidence from the Emergency Department. *Working Paper, Harvard University (available at https://scholar.harvard.edu/saghafian/publications)*. Published online 2019.