To Batch or Not to Batch: Test-Ordering Variability in the Emergency Department and the Impact on Care Delivery

Jacob Jameson, MS[1]
Soroush Saghafian, PhD[2]
Robert S. Huckman, PhD[3]
Nicole Hodgson, MD[4]

Summary
Emergency Department (ED) patients may receive varying diagnostic workups and dispositions based on physician factors instead of solely based on presenting conditions. This study delves into the contrasting practices of batch-ordering multiple tests simultaneously versus the sequential ordering of tests based on previous results. Our analysis revealed stark differences in physician diagnostic approaches, even when working in similar environments. Findings suggest that physicians who predominantly make use of batching ("batchers") tend to order more tests, which is associated with longer lengths of stay and increased costs. In contrast, other physicians ("non-batchers") order fewer tests, which is associated with lower lengths of stay and costs, without any impact on primary ED outcome measures, such as the 72-hour rate of return. Thus, our results suggest an "information gain" advantage in the non-batching strategy: by ordering sequentially, non-batchers obtain the diagnostic information needed with a lower number of tests, enabling them to deliver the same quality of care more efficiently (e.g., with a lower length of stay and cost) than batchers. Finally, our study shows that the decision to batch order diagnostic tests can be optimized for each patient using a few variables, including acuity, chief complaints, and the ED volume at arrival.

[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

Introduction

Emergency Departments (EDs) sit at the forefront of healthcare delivery, often grappling with inefficiencies that impact patient outcomes and the broader health system. One potential determinant of these inefficiencies lies in the variability in physicians' approaches to diagnostic test ordering. Drawing from data from over 43,000 ED visits at the Mayo Clinic of Arizona, we observed pronounced variability in physicians' tendencies to batch-order or sequentially-order diagnostic tests. This variation persisted even among physicians practicing under the same guidelines within the same hospital setting, serving patients randomly assigned to them.

The data indicated that differences in test-ordering practices were not mere anomalies but were rooted in individual physician decisions and systematic across complaint categories (Figure 1). Notably, physicians who batch (do not batch) in one category of complaints tend to also batch (not batch) in other categories. Our preliminary findings underscore a potential imperative: designing more effective and efficient care delivery methods by refining guidelines that target test-ordering strategies. This is crucial in optimizing ED operations[1,2], curbing unnecessary expenditures, and enhancing patient outcomes.

Study Data and Methods

Setting and Sample: Our data is from the Emergency Department at the Mayo Clinic of Arizona, a tertiary care hospital. We performed a retrospective review of ED operational data from 10/6/2018 through 12/31/2019. During our study's timeframe, the ED recorded approximately 43,000 visits. The department is staffed only by board-eligible or board-certified emergency physicians, not nurse practitioners or physician assistants. Of note, residents in rotation oversaw a low fraction, roughly 10%, of the patient volume.

Figure 1

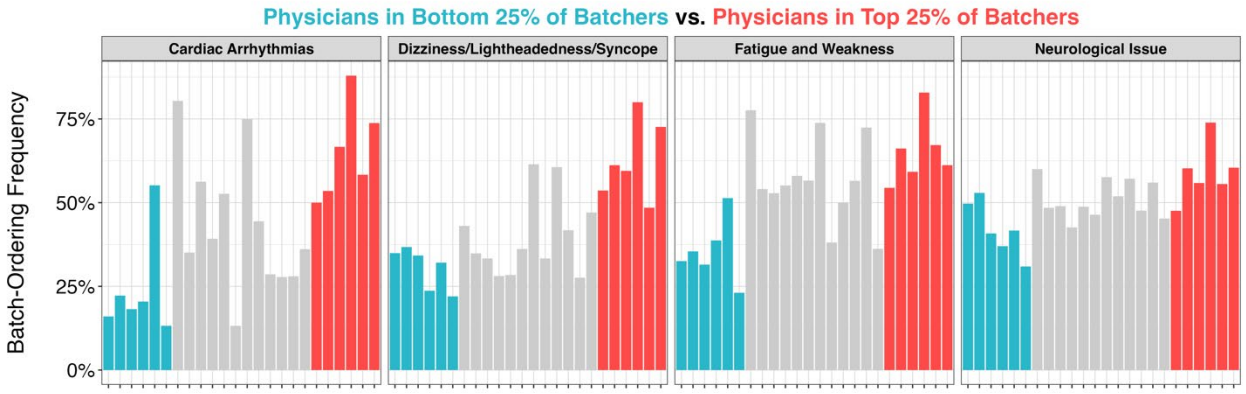

Figure 1 illuminates the marked differences among physicians in their propensity to batch-order diagnostic tests. Physicians are mapped on the x-axis, revealing those with a systematically heightened tendency to batch (in red) compared to their peers who batch less frequently (in blue). The quartile of batchers is calculated based on the batching rate across all complaint areas per physician. Complaints presented in this figure are from patient encounters that exhibit the highest variance in physician batching. Physicians who batch (not batch) in one category of complaints tend to also batch (not batch) in other categories.

Data Compilation: We retrospectively collected comprehensive patient data, encompassing demographics, chief complaints, vital signs, emergency severity index, length of stay, and resource utilization metrics.

Definitions: We designate physicians in our dataset to be of two types: "batchers" and "non-batchers," which we define in the following way:

- *Batcher*: A physician that batch-orders diagnostic tests more than the median physician in the ED across all patient encounters.

- *Non-Batcher*: A physician who batch-orders diagnostic tests less than the median physician in the ED across all patient encounters.

In the context of diagnostic test ordering, we defined batching as placing multiple test orders within 5 minutes of one another. Sensitivity analyses on this cutoff point showed that our results are robust to this definition. Our study differentiated between two distinct batching strategies:

- *Lab + Image Batch*: The batch order for a patient consisted of two or more distinct diagnostic tests with at least one imaging (contrasted CT scan, non-contrasted CT scan, X-ray, ultrasound) test and one lab test.

- *Image + Image Batch:* The batch order for a patient consisted of two or more distinct imaging (contrasted CT scan, non-contrasted CT scan, X-ray, ultrasound) tests.

The distinction is necessary because laboratory tests may be processed in parallel with imaging tests, which have a much more significant effect on patient waiting time. However, a batch of multiple imaging tests requires that the patient is present for each test. Thus, waiting times are more likely to be impacted. Given the varied complexity of patient encounters, the specific tests that make up a batch may provide insight into how the physician trades off between waiting time and uncertainty.

Figure 2 shows the rate at which each batching type occurs for the ten most common chief complaints.

Physician Assignment Process: 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 rotational patient assignment algorithm[3], which practically removes potential selection bias concerns from our analyses. In essence, barring arrival time and shift-level variation, the physician-to-patient matching can be deemed random. Table 1 displays the balance test results, which show that the main complaint and severity for patient encounters are indeed equitably distributed across batchers and non-batchers in our study's cohort. In other words, due to the random assignment, both types of physicians served the same portfolio of patients. Thus, the differences in their test ordering behavior are not due to differences in the complexities of patient conditions. Batchers and non-batchers also worked during comparable times of day and months of the year, making their shift assignment irrelevant.

Figure 2

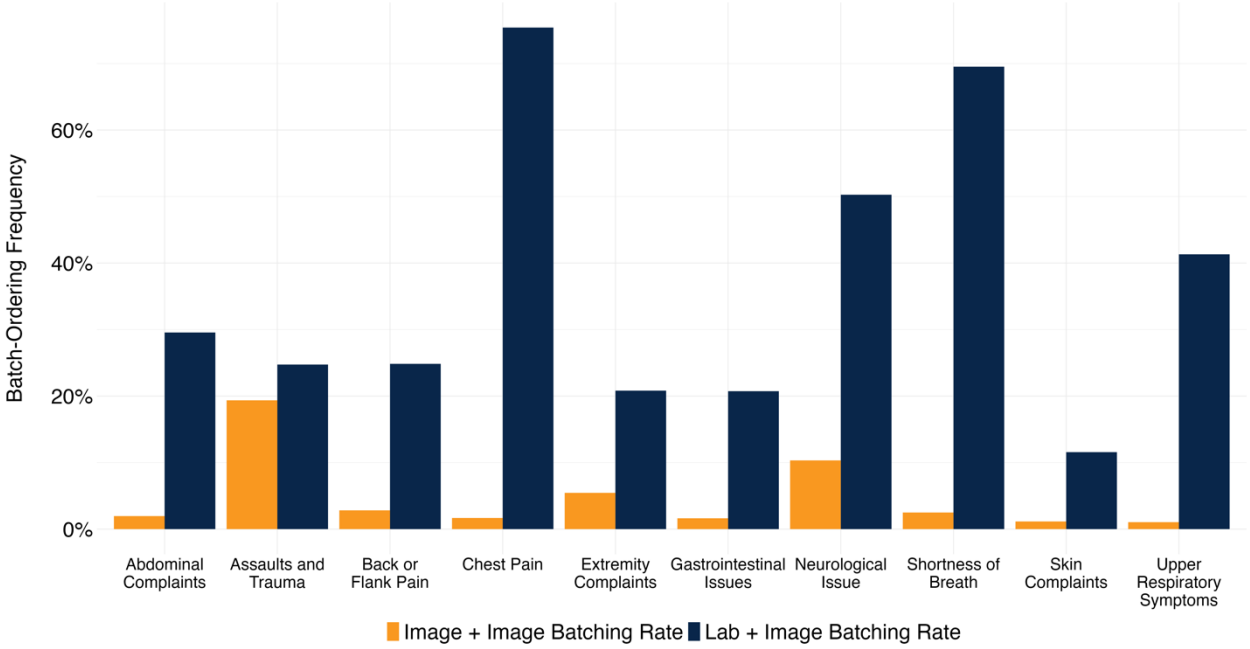

Figure 2 displays two statistics for the ten most common major chief complaint categories observed in our ED visit sample: the unadjusted batching rate for lab-image batching and image-image batching.

Study Results

Table 2 provides a comprehensive breakdown of the patient encounters in which a batcher or non-batcher saw patients during an ED encounter from our baseline sample.

Differences were noted between the two cohorts. Patients seen by a batcher tended to have a longer ED length of stay ($p<0.001$); however, the variability of the length of stay was much shorter (SD of 150 vs. SD of 460).

Patients seen by batchers were also more likely to have more diagnostic tests ordered for them ($p<0.001$). Across all diagnostic tests, an order was more likely to be placed when a batcher saw the patient. This difference was stark for X-ray orders, where a batcher ordered an X-ray for 52% of patient encounters, and a non-batcher ordered an X-ray for 43% of patient encounters ($p<0.001$). This variability in the frequency and types of tests ordered is notable considering the balanced patient panels we observe in Table 1, indicating the need for concrete measures of patient complexity and severity[4] to inform a physician's decision to batch-order diagnostic tests. Interestingly, the likelihood of a 72-hour return was comparable between groups, indicating that the extra tests ordered by batchers did not influence a primary outcome measure used to approximate delivery quality in emergency medicine. Put simply, these results suggest an "information gain" advantage: by ordering sequentially, non-batchers could obtain the diagnostic information needed to serve their patients with a lower number of tests and lower length of stay, delivering the same quality of care more efficiently than batchers.

Upon further examination of the data in Table 2, when limited to encounters where two or more tests were ordered, we found that the ED length of stay for batchers and non-batchers was comparable ($p=0.16$). This finding indicates that the main "information gain" yielding efficiency

in care delivery is obtained when a non-batcher orders a single test and finds the information obtained sufficient to avoid further tests.

Table 1

Chief Complaints	**F-Statistic**	***Pr(> F)***
Abdominal Complaints	2.587	0.108
Back or Flank Pain	1.637	0.201
Chest Pain	0.407	0.524
Extremity Complaints	1.847	0.174
Falls, Motor Vehicle Crashes, Assaults, and Trauma	0.023	0.880
Gastrointestinal Issues	0.105	0.746
Neurological Issue	0.135	0.713
Shortness of Breath	1.324	0.250
Skin Complaints	0.383	0.536
Upper Respiratory Symptoms	0.017	0.896
Emergency Severity Index (ESI)	**F-Statistic**	***Pr(> F)***
ESI 1 or 2	0.011	0.915
ESI 3, 4, or 5	0.010	0.921
Vital Signs	**F-Statistic**	***Pr(> F)***
Tachycardic	0.118	0.731
Tachypneic	0.043	0.836
Febrile	0.936	0.333
Hypotensive	1.127	0.288

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.

Discussion

ED physicians inherently seek prompt and precise diagnoses. Our findings suggest that the nature of this diagnostic process may be influenced by physician ordering practices in addition to the intricacies of individual patient conditions[5]. When a physician orders and evaluates one test, they gain insights into the patient's condition, which can then inform whether further tests are necessary. If the subsequent test is deemed unnecessary based on the initial test's results, then the sequential method offers a more efficient approach than batch ordering. Our findings suggest that non-batchers, who might be employing this "information gain" strategy, seem to manage patient care more efficiently, as evidenced by lower ED lengths of stay and comparable 72-hour return rates, despite ordering fewer tests.

Given the impact of physician variation on outcomes such as length of stay, expenditures, and readmission risk, it is crucial to gain insights into which approach—batching or sequencing—is superior based on a patient's condition. Figure 3 underscores this by showing a clear trend: physicians who prefer batch ordering often order more tests. This practice has implications: an

extended length of stay and a surge in associated expenditures from potentially superfluous tests. Drawing on insights from related literature[6], we determined that the financial burden of these extra tests is substantial. However, determining the optimal strategy that balances concerns related to costs and quality requires a deeper understanding of the precise inefficiency at play.

Table 2

Variable	**Overall**, N = 43,328[1]	**Batcher**, N = 17,861[1]	**Non-Batcher**, N = 25,467[1]	**p-value**[2]
Disposition				
Admit	8,665 (20%)	3,945 (22%)	4,720 (19%)	<0.001
Discharge	28,191 (65%)	11,228 (63%)	16,963 (67%)	
Other	6,472 (15%)	2,688 (15%)	3,784 (15%)	
Number of Tests Ordered				
0	4,801 (11%)	1,508 (8.4%)	3,293 (13%)	<0.001
1	11,828 (27%)	4,432 (25%)	7,396 (29%)	
2	17,959 (41%)	7,816 (44%)	10,143 (40%)	
3	7,356 (17%)	3,470 (19%)	3,886 (15%)	
4	1,292 (3.0%)	599 (3.4%)	693 (2.7%)	
5	92 (0.2%)	36 (0.2%)	56 (0.2%)	
ED LOS (min)				<0.001
Mean (SD)	251 (365)	263 (150)	243 (460)	
72 Hour Return	1,569 (3.6%)	626 (3.5%)	943 (3.7%)	0.29
Ultrasound Ordered	4,940 (11%)	2,180 (12%)	2,760 (11%)	<0.001
Non-Contrast CT Ordered	8,978 (21%)	4,007 (22%)	4,971 (20%)	<0.001
Contrast CT Ordered	8,155 (19%)	3,460 (19%)	4,695 (18%)	<0.001
Laboratory Tests Ordered	33,183 (77%)	14,177 (79%)	19,006 (75%)	<0.001
X-ray Ordered	20,186 (47%)	9,226 (52%)	10,960 (43%)	<0.001

[1] *n(%)*

[2] Pearson's Chi-squared test; Welch Two Sample t-test

Figure 3

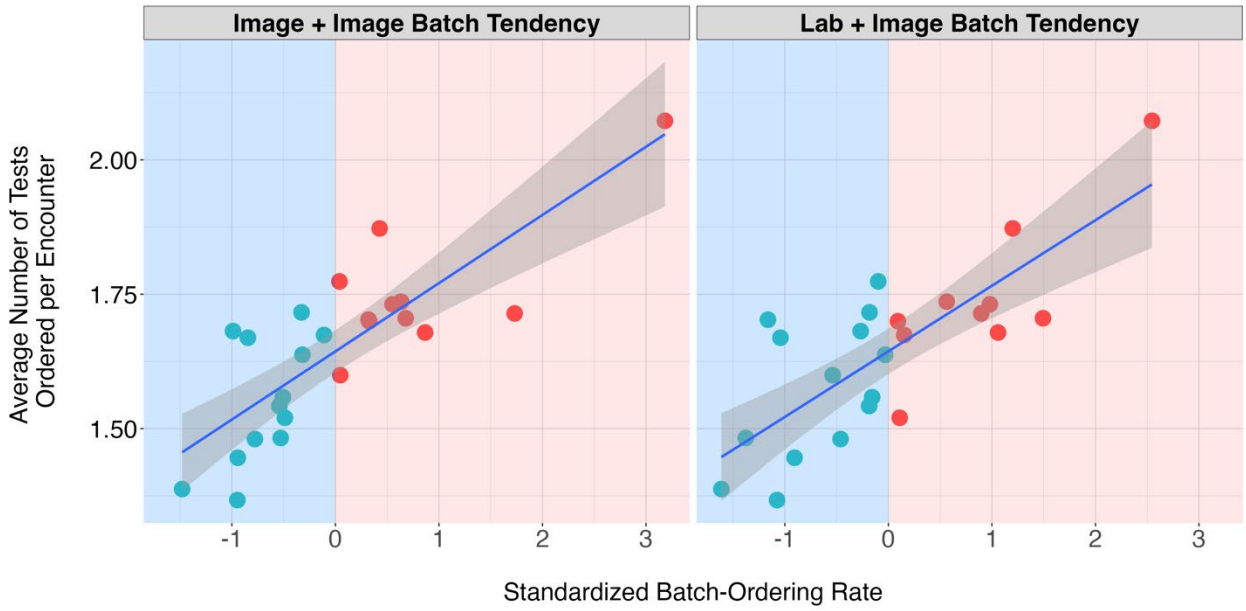

Figure 3 displays the relationship between the standardized batch-ordering rate and the average number of tests ordered per encounter for physicians within our study sample. Notably, physicians with a standardized batch-ordering rate above 0 (i.e., a batch rate greater than the average of the sample) tend to order more diagnostic tests on average as compared to physicians with a standardized batch-ordering rate below 0 (i.e., a batch rate below the average of the sample)

Notably, our results raise the question of whether—if presented with a choice upon entering the ED—a given patient would be better served by a batcher or a non-batcher[7]. Regarding the preferred patient-physician assignment, the question of who should see the patient has received some attention in recent years in different settings[8]. However, this preferred assignment has not been explored in the ED, where diagnostic testing is critical. Our results indicate that given the positive overall association between a physician's tendency to batch and order more tests, being seen by a batcher on average may lead to efficiency losses. However, as shown in Figure 4, this result is nuanced by specific patient and hospital-level conditions and the particular objective being optimized for.

Figure 4 displays the results of our findings based on the patient's Emergency Severity Index (ESI) and the ED's current occupancy, stratified by four of the most common patient complaints. Results indicate that, except for the most critical cases, being seen by a non-batcher is generally preferred when the goal is to reduce the length of stay and the total number of diagnostic tests ordered. Interestingly, constrained ED capacity when the patient is served appears to play less of a role in determining the preferred mode of ordering tests. Nevertheless, there is apparent heterogeneity in optimal testing strategy by acuity, complaint, and the physician/hospital objective (i.e., decreased length of stay vs. decreased likelihood of 72-hor return). Given the importance of optimizing ED operations and patient outcomes considering tradeoffs in information gain, speed, and quality[9], we hope our findings motivate future research into determining when a physician should batch order tests.

Conclusion

The diagnostic process is central to ED operations and impacts patient outcomes. Our investigation into diagnostic test ordering patterns reveals stark differences, even among physicians operating within identical environments and adhering to similar guidelines.

Central to our findings is the batch-ordering practice, adopted in approximately 39% of diagnostic procedures within our study's purview. Our data indicates that not batching ED tests yields an information gain advantage, enabling physicians to reduce the number of tests and shorten the patient length of stay without any negative impact on patient outcomes such as 72-hour rate of return. Future studies, however, are required to validate our findings and obtain more insights into optimal testing strategies.

Figure 4

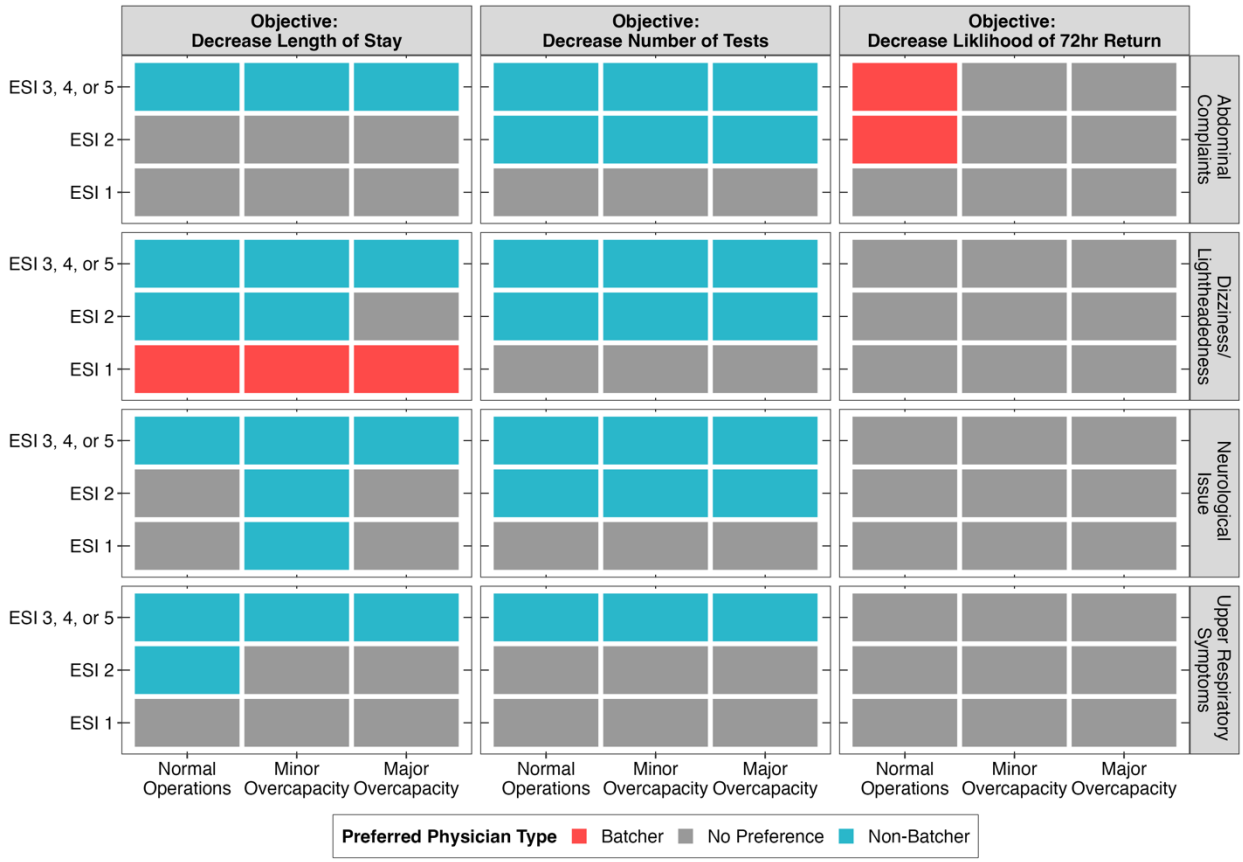

Figure 4 illustrates the interplay between testing strategy preference for various emergency severity, complaint, and ED volume levels. On the x-axis, emergency department volume levels are displayed, while the y-axis conveys the 'Emergency Severity Index.' The color gradient differentiates the preferred physician given the objective: Batcher in red and Non-Batcher in blue. Faceted by three crucial outcomes – the number of tests, length of stay, and 72-hour return rates – the visualization underscores the heterogeneity in the optimal testing strategy. Preferred Physician Type was determined based on the average marginal effects of a linear regression model stratified by the complaint.

Our findings raise the question of whether the chosen diagnostic approach optimally serves patients from both a quality and cost perspective. The derived insights suggest a need for revisiting and potentially revising test ordering guidelines. Such revisions, grounded in

efficiency and patient welfare, could streamline ED operations and extend current approaches to improving them[1,2,4], curtailing healthcare expenditures and amplifying the quality of patient care.

The variation we observe across physicians could stem from myriad sources, including physicians' training, accumulated experience, or ingrained diagnostic philosophies[10]. These influences could drive a physician towards a particular testing methodology. Efficient, optimized diagnostic methods that could be communicated to physicians (e.g., via training programs) hold the promise of reducing healthcare expenditures and enhancing patient care quality.

As research in this area of inquiry progresses, moving beyond associative insights is imperative. While we have taken advantage of the random assignment in our setting, future research should continue to explore causality, dissecting the intricate relationship between batch ordering, potential over-testing, and their collective impact on patient outcomes. A diagnostic landscape that prioritizes patient well-being through this comprehensive understanding could be influential in improving care delivery.

References

1. Saghafian, S., Hopp, W.J., Van Oyen, M.P., Desmond, J.S., & Kronick, S.L. (2012). "Patient Streaming as a Mechanism for Improving Responsiveness in Emergency Departments." Operations Research, 60(5), 1080-1097.

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

3. Traub, S.J., Stewart, C., Didehban, R., Bartley, A., Saghafian, S., Smith, V., Silvers, S., LeCheminant, R., & Lipinski, C.A. (2016). "Emergency Department Rotational Patient Assignment." Annals of Emergency Medicine, 67(2), 206-215.

4. Saghafian, S., Hopp, W.J., Van Oyen, M.P., Desmond, J.S., & Kronick, S.L. (2014). "Complexity-augmented triage: A tool for improving patient safety and operational efficiency." Manufacturing & Service Operations Management, 16(3), 329-345.

5. Hodgson, Nicole R et al. "Are testers also admitters? Comparing emergency physician resource utilization and admitting practices." *The American Journal of Emergency Medicine* vol. 36,10 (2018): 1865-1869. doi:10.1016/j.ajem.2018.07.041

6. Müskens, J.L.J.M., Kool, R.B., van Dulmen, S.A., et al. (2022). "Overuse of diagnostic testing in healthcare: a systematic review." BMJ Quality & Safety, 31, 54-63.

7. Maria R. Ibanez, Jonathan R. Clark, Robert S. Huckman, Bradley R. Staats (2017) Discretionary Task Ordering: Queue Management in Radiological Services. Management Science 64(9):4389-4407.

8. Atkinson, M.K., & Saghafian, S. (2023). "Who should see the patient? On deviations from preferred patient-provider assignments in hospitals." Health Care Management Science, 26, 165–199.

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

10. Abaluck, J., Agha, L., Kabrhel, C., Raja, A., & Venkatesh, A. (2016). "The Determinants of Productivity in Medical Testing: Intensity and Allocation of Care." American Economic Review, 106(12), 3730-64.