A Comparison of Traditional & Modified Fuzzy Delphi in Identification of Metrics for ED Performance Associated with Medical Surges

Egbe-Etu Etu, Leslie Monplaisir, *Ph.D.*, Celestine Aguwa, *Ph.D.*, Sara Masoud, *Ph.D.*
Department of Industrial & Systems Engineering
Wayne State University
Detroit, MI 48202, USA
fw7443@wayne.edu; leslie.monplaisir@wayne.edu; bd4891@wayne.edu;
saramasoud@wayne.edu

Suzan Arslanturk, *Ph.D.*
Department of Computer Science
Wayne State University
Detroit, MI 48202, USA
Suzan.arslanturk@wayne.edu

Ihor Markevych
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213, USA
imarkevy@andrew.cmu.edu

Joseph Miller, MD
Departments of Emergency Medicine and Internal Medicine
Henry Ford Hospital
Detroit, MI 48202, USA
Jmiller6@hfhs.org

Abstract

The COVID-19 pandemic, as an instance of medical surge, has presented significant challenges to healthcare systems around the world. Efficient operations and management have been a major challenge for hospital administrators during this pandemic. The objective of this research is to conduct a comparative study of two fuzzy Delphi approaches, the traditional fuzzy Delphi (TFD) and modified fuzzy Delphi (MFD), for metrics identification in a hospital surge scenario. To the best of our knowledge, literature lacks a comparison between decision outcomes obtained based on the TFD and other proposed fuzzy Delphi methods. Questionnaire data from our prior work is used for this comparative analysis. The results illustrate that twenty-one metrics based on the TFD method and twenty metrics based on the MFD method were identified for measuring ED's efficiency. The stability of the group responses was confirmed using the U test ($p = 0.0001$ and 0.0047) for TFD and MFD, respectively. The analysis shows that the MFD method outperformed the TFD approach in handling the fuzziness in experts' responses.

Keywords
COVID-19, Performance Measurement, ED metrics, Surge, Fuzzy Delphi

1. Introduction
Although hospital emergency departments (ED) always play an essential role in delivering health care to patients, their role is crucially highlighted during public health emergencies such as a disease outbreak and natural or man-

made disasters. EDs, as the vanguards of healthcare facilities, have been significantly overwhelmed by the influx of patients during the recent Coronavirus outbreak, which in return has impacted their performance and efficiency noticeably. For instance, in March 2020, the intensive care units (ICU) in Lombardi, Italy, experienced an exponential increase in patients which affected the hospital's performance (Grasselli, Pesenti, and Cecconi 2020). Also, in the US, states such as Texas, Arizona, and Florida are experiencing an influx of patients causing ICUs to exceed capacity (Crist 2020). As a result, hospital resources (i.e., equipment and personnel) are over-utilized or unavailable due to the increasing number of infected patients seeking care and the quarantined and deceased healthcare workforce. With the reduction in frontline medical staff and equipment, the ability to cater to disease-ridden patients' health needs is significantly affected.

As the healthcare system struggles to sustain the uncontrolled outbreak, a more robust containment measure is the only realistic option to avoid a total collapse. To achieve this, hospital administrators must adopt a performance management strategy to optimize healthcare resources and improve the quality of care rendered to patients in the event of a medical surge. As multiple metrics control the operation and performance of the EDs, many hospital administrators encounter challenges in planning and allocating resources to prevent waste and inefficiencies while improving the quality of care when responding to disasters. Currently, most EDs are struggling and do not have the right means or resources to cater to the influx of infected patients caused by the pandemic; investing in the correct equipment given the limited resources can potentially lessen the burden. Literature shows different techniques that can be used to measure the performance of hospital EDs. The methods include parametric techniques that estimate frontier production functions and calculate technical efficiencies of the given departments (Rezapour et al. 2019). The stochastic production frontier is an example of a parametric method that can be applied to estimate hospitals' capacity and capacity through a modification of the inputs incorporated in the production frontier (Rezaei et al. 2016; Rosko and Mutter 2008). Also, non-parametric techniques such as data envelopment analysis are employed for measuring the relative efficiency of hospitals (Chang 2004). For the techniques mentioned above to work, there needs to be a defined set of metrics to serve as input and output variables for measuring the efficiency of EDs.

Presently, there is a minute amount of research on relevant metrics for assessing the ED's performance during a surge event. Hospital administrators have found it difficult to comprehensively evaluate the EDs or ascertain metrics that should be prioritized for improving operations and processes while confronting disasters. Núñez et al. (2018) investigated different performance indicators in Chilean EDs but focused on the regular operating conditions. Similarly, (Khalifa and Zabani 2016) developed a three-step process for identifying the significant variables related to ED performance. The study focuses on regular operating conditions. Madsen et al. (2016) studied different hospital quality indicators using a modified-Delphi process. Personal biases from the steering committee was a significant limitation of the study. Rezapour et al. (2019) developed a three-round Delphi process to identify variables for measuring public hospitals' efficiency in Iran. The study's weakness is in the subjectivity and vagueness of panel member responses. In other areas, (Chang, Hsu, and Chang 2011) established a fuzzy Delphi model for selecting appropriate hydrogen production technology in Taiwan. Kuo and Chen (2008) applied the fuzzy Delphi method to develop key performance appraisal indicators for the service industries. Zarei et al. (2019) applied the fuzzy Delphi technique to reach a consensus on Iran's drug allocation measures to improve the process and deal with drug scarcity.

Much research has been done on variable identification and efficiency measurement of hospitals or EDs during normal operating conditions. However, the literature lacks a comparison between decision outcomes obtained from traditional fuzzy Delphi and other proposed fuzzy Delphi methods (Saffie and Rasmani 2016). This paper seeks to conduct a comprehensive study of two fuzzy Delphi approaches, TFD and MFD, for metrics identification in a hospital surge scenario. This is a growing area of concern due to the current scale of the Coronavirus pandemic and the strain on hospital resources.

2. Methods

In this section, we discuss the two methods used in this comparative analysis: the TFD and MFD methods. The TFD method was first developed by (Murray, Pipino, and van Gigch 1985) to address some of the drawbacks of the Delphi technique. It is an analytical method based on the combination of probabilistic fuzzy theories and human linguistic variables in surveys while making-decisions (Hsu, Lee, and Kreng 2010). The MFD method proposed by (Etu et al. 2020), as the latest version of imposed improvement on TFD, is a three-stage process optimized by the consistency aggregation method (CAM) for the selection, identification, stability determination, and ranking of any given variable based on expert responses. MFD adheres to the characteristics of the Delphi procedure and explains the controlled

feedback for the iterative process. To fill the identified gap in the literature, we perform a comparative analysis of the two mentioned methods in subsequent paragraphs.

Based on the data provided by (Etu et al. 2020), the identified metrics in literature are grouped into Capacity, Temporal, Quality, Outcomes, and Financial expenditures. These metrics are used to survey the medical professionals via questionnaires to summarize the impact of the provided set of metrics on the ED's performance during a surge event. The participants (e.g., hospital administrators, medical directors, senior ED physicians, resident physicians, registered nurses) are experts in hospital ED management, efficiency analysis, and continuous quality improvement. To ensure the quality of the process, we applied the same dataset for both methods. The utilized questionnaire data is obtained from (Etu et al. 2020), enabling us to perform the comparative analysis on TFD and MFD methods as displayed in Figure 1.

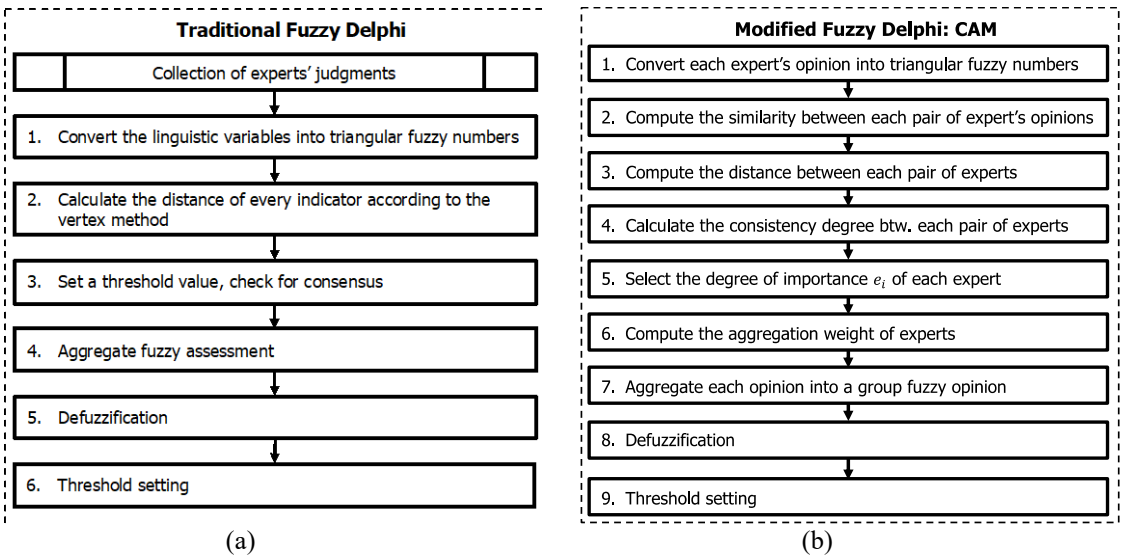

Figure 1. (a) Flowchart of the TFD method and (b) Flowchart of the MFD method (Etu et al. 2020)

2.1 Traditional Fuzzy Delphi (TFD) method

The TFD method has been extensively described in the literature by (Chang, Hsu, and Chang 2011; Hsu, Lee, and Kreng 2010; Kuo and Chen 2008; Ho and Wang 2008). The TFD method follows a six-step process, which is displayed in Figure 1a. The mathematical expression of the TFD method is as follows: Assume that K medical experts are invited to evaluate the importance of any given healthcare performance elements, i and the associated metrics, j using linguistic variables (i.e., strongly disagree to strongly agree). We represent the linguistic variables as a set of triangular fuzzy numbers on the interval [0 – 1]. Let the fuzzy numbers $\tilde{r}_{ij}^{k}$ be the rating of elements i with respect to metrics j and $\tilde{w}_j^k$ be the jth metric weight of the kth expert for $i = 1, \dots, m; j = 1, \dots, n; k = 1, \dots K$ and $\tilde{r}_{ij} = \frac{1}{k}\left[\tilde{r}_{ij}^{1} \oplus \tilde{r}_{ij}^{2} \oplus \dots \oplus \tilde{r}_{ij}^{K}\right]$; $\tilde{w}_j = \frac{1}{k}[\tilde{w}_j^1 \oplus \tilde{w}_j^2 \oplus \dots \oplus \tilde{w}_j^K]$. For each medical expert, we use the weighted Euclidean distance to compute the distance between the average $\tilde{r}_{ij}$ and $\tilde{r}_{ij}^{K}$ and the distance between the average $\tilde{w}_j$ and $\tilde{w}_j^k$, $k = 1, \dots K$. The distance between two fuzzy numbers $\tilde{m} = (m_1, m_2, m_3)$ and $\tilde{n} = (n_1, n_2, n_3)$ is computed by $d(\tilde{m}, \tilde{n}) = \sqrt{\frac{\sum_{i=1}^{n}(\tilde{m}_i - \tilde{n}_i)^2}{3}}$. We set a threshold value to check for consensus if the medical expert's evaluation data is less than the threshold value of 0.2 or the percentage of achieving a group consensus is greater than 60%, then we go to step 5; otherwise, a new round of survey is required. We collect and aggregate each expert opinion to obtain a group opinion, and the centroid defuzzification approach is used to convert the fuzzy values into crisp values. We test the stability of the expert responses using multivariate analysis and rank the metrics based on importance.

2.2 Modified Fuzzy Delphi (MFD) method

This section describes the MFD method developed by (Etu et al. 2020) for identifying critical metrics that affect the performance of the ED during a surge scenario. The method involves a nine-step process illustrated in Figure 1b. The

mathematical expression of the MFD method is as follows: Given that $O = \{O_n | n = \overline{1, N}\}$ is a set of N healthcare performance elements, where each element denotes a limited set of metrics (i.e., P_n), $I_n = \{I_{np} | p = \overline{1, P_n}\} \forall O_n \in O$. We form a panel of medical professionals, M of size q to analyze each metric, $M_{np} = \{M_{npq} | q = \overline{1, Q_{np}}\} \forall I_{np} \in I_n$. In view of their skills, every professional uses the Likert scale to assess the given questionnaire. Anonymity amongst experts is vital to prevent exchange of information, a random distribution of opinions for each metric is the most probable. All responses are on a Likert scale (1 – 5 level scale), which is converted to triangular fuzzy numbers ranging from 0 to 1. We collect and aggregate each expert opinion to obtain a group opinion, and a threshold for each metric is obtained. Any metric that exceeds the agreement level is accepted, while those below the agreement level is modified and a new survey is sent out. We use the multivariate analysis to ensure the expert responses' stability during the iterative process and then rank each metric using a permutation operation.

The two methods are implemented in the Python environment, and the source codes for each algorithm are provided in the links – MFD method (Etu et al. 2020): https://github.com/Dergel0806/Consistency-Aggregation-Method. TFD method: https://github.com/Dergel0806/Traditional-Delphi-Method

3. Results

After researching the selected article database, twenty-nine potential metrics were identified and extracted. We clustered the metrics into five elements which include twelve capacity-related metrics {*ED beds, Intensive care unit (ICU) beds, Physician staffing, Midlevel provider staffing, Nurse staffing, Patient acuity level, Physician staffing per patient seen, Nurse staffing per patient seen, Backup physician, Backup nurse, Patient care compromised, and medical support personnel*}. Eight temporal metrics {*High acuity, Low acuity, Time to start of treatment, Admit ED length of stay (LOS), Time to triage, Discharge ED LOS, and Time to treatment condition*} and three quality-related metrics {*Employee fatigue, Employee satisfaction, and Medical errors*}. Two outcome-related metrics {*Patients hospitalized and patient transfers*}, and four financial metrics {*Increase diagnostic test, Increase ED treatment, Increase ED revenue, and Increase in non-labor cost*}. A description of each metric is provided in (Etu et al. 2020). The identified metrics were selected based on the availability of data, measurability, and ease of implementation in the ED.

The questionnaire dataset contained healthcare experts' responses over a two-round iterative process; the first-round survey data consisted of 45 expert responses and were collected in December 2019 (Etu et al. 2020). The second-round data was collected in February 2020 and consists of 23 expert responses (Etu et al. 2020). Table 1 shows the profile of the panel members for the first and second rounds.

Table 1. Panel members' characteristics (Etu et al. 2020)

Characteristics	First Round	Second Round
	Number (percentage of frequency)	Number (percentage of frequency)
Gender		
Female	22 (48.89)	12 (26.67)
Male	23 (51.11)	11 (24.44)
Years of Experience with ED processes & operations during a surge		
< 1 year	0	0
1 – 3 years	8 (17.78)	5 (21.74)
4 – 6 years	6 (13.33)	4 (17.39)
> 7 years	31 (68.89)	14 (60.87)
Degree		
Associate degree	3 (6.67)	0
Bachelor's degree	12 (26.67)	5 (11.11)
Ph.D./MD	30 (66.67)	18 (40.00)
Types of Stakeholders		
Physicians	17 (37.78)	11 (47.82)
Resident physicians	6 (13.33)	2 (8.69)
Registered / Clinical nurses	13 (28.89)	3 (13.04)
Midlevel providers	3 (6.67)	3 (13.04)
Medical directors	2 (4.44)	2 (8.69)

Administrators	3 (6.67)	1 (4.35)
Paramedic	1 (2.22)	1 (4.35)

Eighty-two percent of the respondents, which include physicians, resident physicians, registered nurses, mid-level providers, medical directors, administrators, and paramedic work in the ED while the remaining eighteen percent of respondents work in other hospital departments but have vast experience in ED operations management during a surge event. The panel members have a minimum of 3 years' experience, a median of 7 years' experience and a maximum of 11 years' experience working in the ED during surge-related scenarios. The combined total ED-related experience for the respondents is 304 years. We use this information to calculate a percentage of the weighting of experience for each expert. The findings from implementing the TFD and MFD methods in the first round are illustrated in Table 2.

Table 2. First-round results for TFD & MFD methods

Elements	Metrics	TFD		MFD	
		Avg. of fuzzy numbers	Consensus (threshold > 64)	Avg. of fuzzy numbers	Consensus (threshold > 53)
Capacity	ED beds	0.128	95.556	0.113	77.879
	Intensive care unit (ICU) beds	0.152	91.111	0.167	75.063
	Physician staffing	0.33	35.556	0.307	49.958
	Midlevel provider staffing	0.348	46.667	0.328	50.148
	Nurse staffing	0.218	86.667	0.167	62.184
	Patient acuity level	0.546	51.111	0.586	47.281
	Physician staffing per patient seen	0.594	44.444	0.647	47.618
	Nurse staffing per patient seen	0.647	80	0.709	52.943
	Backup physician	0.486	62.222	0.464	48.229
	Backup nurse	0.549	57.778	0.56	45.711
	Patient care compromised	0.628	80	0.668	51.371
	Medical support personnel	0.353	40	0.322	48.941
Temporal	High acuity	0.377	37.778	0.35	45.114
	Low acuity < 60 mins	0.168	86.667	0.126	69.32
	Admit ED LOS < 6 hrs	0.139	88.889	0.099	75.033
	Discharge ED LOS < 4 hrs	0.153	86.667	0.129	72.361
	Time to triage	0.353	62.222	0.334	53.917
	Time to start of treatment	0.319	53.333	0.288	52.762
	Time to ED bed	0.324	48.889	0.283	50.826
	Time to treatment condition	0.299	66.667	0.268	57.95
Quality	Employee fatigue	0.724	100	0.741	60.947
	Employee satisfaction	0.14	91.111	0.119	75.279
	Medical errors	0.597	64.444	0.602	52.572
Outcomes	Patients hospitalized	0.458	93.333	0.454	64.093
	Patient transfers	0.353	35.556	0.322	49.14
Financial expenditures	Increase diagnostic test	0.538	64.444	0.54	49.038
	Increase ED treatment	0.476	48.889	0.46	44.28
	Increase ED revenue	0.387	68.889	0.385	61.557
	Increase in non-labor cost	0.564	66.667	0.571	51.65

*The highlighted values show consensus was reached based on group opinions for each metric
*Note – ED: Emergency Department; LOS: Length of stay; TFD: Traditional Fuzzy Delphi; MFD: Modified Fuzzy Delphi

The first-round results indicate that only five and three out of twelve capacity metrics achieved a group consensus in TFD and MFD methods, respectively. Feedback from the panel suggests that the provision of a treatment or triage space is essential and would help improve the hospital's ability to provide care to patients when responding to a medical surge. As for the temporal elements, four and five out of eight metrics achieved a group consensus in the TFD and

MFD, respectively. Panel members agreed that poor communication between patients and nurses led to prolonged waiting times. For the quality elements, three and two out of the three suggested metrics achieved a group consensus in this category in TFD and MFD methods, respectively. Qualitative comments in this category show the need to provide quality medical care when patients surpass the available resources to avoid medical errors. Outcome elements: both TFD and MFD methods had one metric achieved consensus, respectively. Feedback from the panel indicates that leaving patients in hallways without adequate treatment affects their outcome. As for financial elements, three and one out of four suggested metrics achieved a group consensus in TFD and MFD methods, respectively. The results of the first-round analysis, mode of consensus, and feedback sorted by each metric is sent to the respondents. The respective thirteen and seventeen metrics of TFD and MFD methods that did not achieve agreement amongst the panel members are used to create a revised survey questionnaire for the second Delphi round for final decision-making and agreement.

The second Delphi round analysis shows that five metrics in the TFD method and eight metrics in the MFD method out of the seventeen metrics achieved consensus and are recognized as suitable metrics for assessing the ED's efficiency when responding to a medical surge. Metrics like physician staffing, backup physician, backup nurse, medical support personnel, high acuity < 30 mins, time to ED bed amongst others did not achieve consensus in the second-round. The health professionals believe that this metrics do not significantly contribute to the performance of the ED during a medical surge. The results and controlled feedback are sent as a report to the respondents who participated in the second-round survey. Table 3 shows the results of the second-round fuzzy Delphi analysis.

Table 3. Second-round results for TFD & MFD methods

Elements	Metrics	TFD		MFD	
		Avg. of fuzzy numbers	Consensus (threshold > 70)	Avg. of fuzzy numbers	Consensus (threshold > 56)
Capacity	Physician staffing	0.478	52.174	0.473	51.15
	Midlevel provider staffing	0.557	69.565	0.571	57.06
	Patient acuity level	0.643	86.957	0.691	56.91
	Physician staffing per patient seen	0.661	91.304	0.665	63.41
	Nurse staffing per patient seen	--	--	0.781	77.34
	Backup physician	0.452	69.565	0.434	54.08
	Backup nurse	0.576	60.87	0.585	49.23
	Patient care compromised	--	--	0.702	56.7
	Medical support personnel	0.557	52.174	0.567	45.87
Temporal	High acuity < 30 mins	0.557	60.87	0.582	48.2
	Time to start of treatment	0.43	78.261	0.43	57.02
	Time to ED bed	0.443	56.522	0.44	43.53
Quality	Medical errors	--	--	0.618	52.25
Outcomes	Patient transfers	0.239	91.304	0.205	65.64
Financial expenditures	Increase diagnostic test	--	--	0.546	52.51
	Increase ED treatment	0.522	78.261	0.537	58.27
	Increase in non-labor cost	--	--	0.579	55.82

*The highlighted values show consensus was reached based on group opinions for each metric
*The spaces with "--" show consensus was reached during the first round of the TFD method

Although the two methods provide quite a similar set of metrics (i.e., 18 metrics in common) as displayed in Table 3. The only exceptions are increase in non-labor cost, medical errors, and increase diagnostic test which achieved consensus in the TFD method while midlevel provider staffing and time to triage achieved consensus in the MFD method. We can infer that using the MFD method caters to the subjectivity of expert opinions in a survey and provides a robust set of metrics for measuring the efficiency of EDs during a surge event.

4. Discussion

Upon merging the results of the two rounds, we obtain a full set of retained metrics (see Figure 2), with interesting patterns. First, we observe from the graph that the TFD consensus is always higher than those of the MFD consensus.

According to our analysis, we discover that the TFD method does not have the similarity measure function like the MFD method, which uses it to determine how close the responses of two healthcare professionals are. The function ensures that the MFD results are not loose, like those of TFD method. Another reason for the scores achieved with the MFD method is the inclusion of expert years of experience as this plays a role in their evaluation. Second, the percentage average of TFD fuzzy numbers is usually lower than the MFD fuzzy numbers, regardless of the first observation. The reason for this lies in the steps taken for the conversion of fuzzy values to crisp values between the two methods. Lastly, we observe that both TFD and MFD consensus follows the same pattern in terms of increase and decrease over different metrics. We believe this is caused by how the metrics are measured and monitored in the ED.

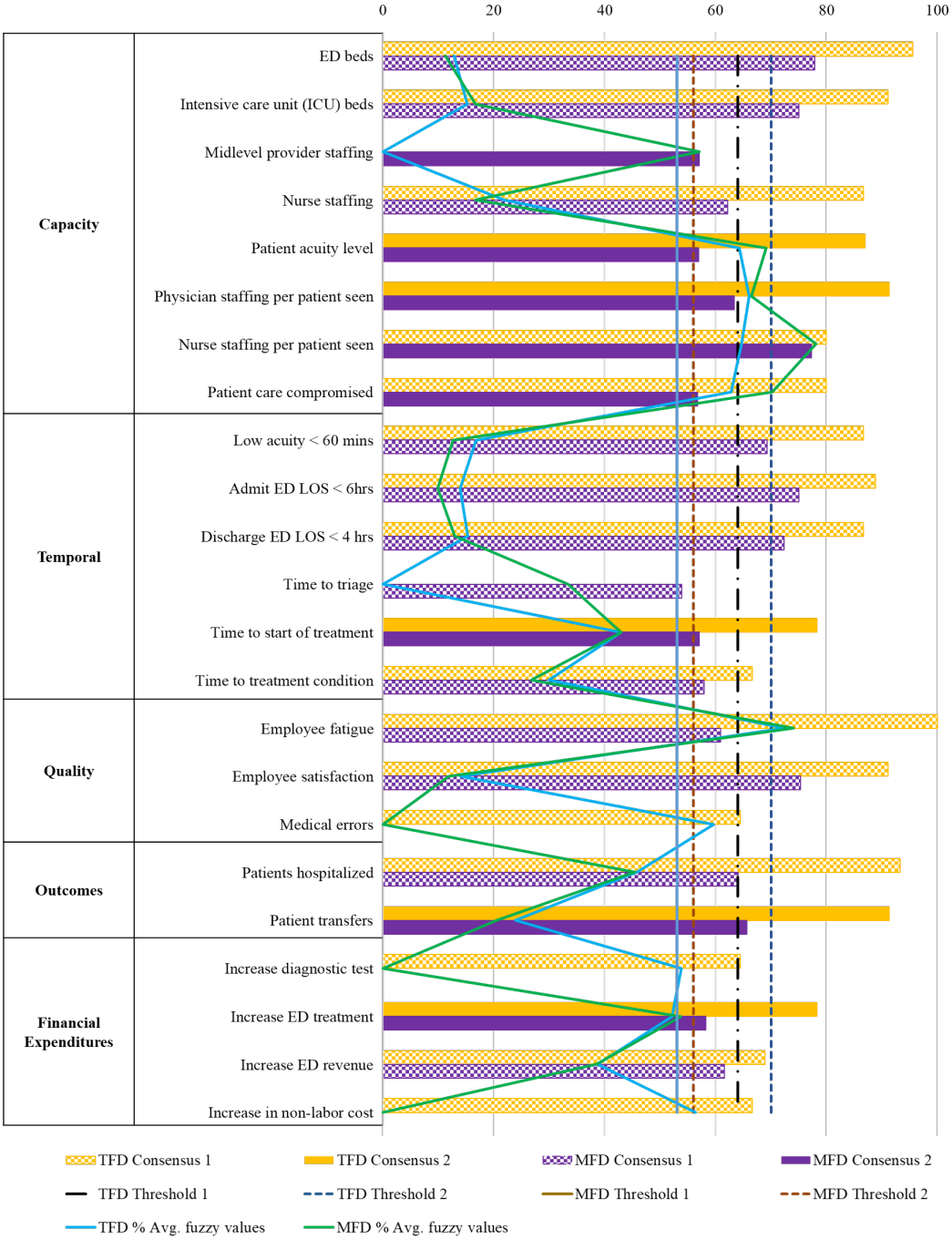

Figure 2: Trends in first & second-round results using TFD and MFD methods

The Delphi procedure is viewed as complete when the agreement and stability levels have been reached between survey rounds and starting a new round will not change the results. A non-parametric means difference test is conducted to confirm stability and convergence of the results. The results obtained with the *U* test shows statistical significance at a $significant\ level\ of\ 0.05\ (z = 62.0\ and\ p = 0.0047)$ for the MFD method. Also, the results obtained with the *U* test for the TFD method shows statistical significance $(z = 10.0\ and\ p = 0.0001)$. However, the MFD results show that changes in the expert responses contributed to a higher agreement level after the first round, as observed in round two. Having achieved stability, the retained metrics are ranked based on importance as specified by the expert panel. Figure 3 presents the final list of metrics based on the importance of ED efficiency measurement during a surge. We can see the difference in retained metrics from the two methods used. A total of twenty-one metrics was retained in the TFD method, while twenty was retained in the MFD method.

Figure 3. Final metrics for ED efficiency measurement

In a theoretical level, the fact that these two methods have 18 metrics in common, while taking different approaches in translating experts' responses, can be counted as a validation of these methods. Although similar in results, the main advantage of the MFD method is proven to be the removal of ties and the clear rankings of metrics. The MFD method achieved that by taking advantage of similarity and distance measures to acquire a consistency index of individual panel member opinion compared to the TFD method that applies only the Euclidean distance measure.

The MFD method ranks ED beds, nurse staffing per patient seen, employee satisfaction, and ICU beds as the most important metrics that affect the ED's performance during a surge while the TFD method ranks employee fatigue, ED beds, patients hospitalized, and physician staffing per patient seen as the most critical metrics. When comparing the rankings, based on expert opinions, ED beds are essential for the treatment of patients during a medical surge than employee fatigue. Without beds, patients cannot get the needed treatment, and this puts a strain on medical personnel, which affects performance. An increase in ED beds requires an increase in nurses or physicians, but in the current COVID-19 pandemic, this is not the case as some healthcare workers are infected, leading to a shortage of frontline medical workers who are stressed and overwhelmed.

For a more practical validation, the performance of the TFD and MFD methods can be evaluated under the ongoing pandemic (i.e., COVID-19). Although the study was not explicitly designed with the COVID-19 in mind, it is holding up with regards to the current events as hospital administrators face challenges with performance and delivering the highest quality service to patients. To validate our findings, we look at the recent coronavirus events happening across three states in the US, namely Florida, Texas, and California. We can see that hospitals in these states are overwhelmed

as the infection rates in those states increase (Walters, Najmabadi, and Platoff 2020; Crist 2020). Hospital performance is greatly affected as administrators struggle to make informed decisions on allocating ED and ICU beds to patients and assigning nurses and physicians to patients with severe illnesses as reported by the TFD and MFD methods. Most hospitals have exceeded capacity, cannot accept patients, and have resorted to transferring patients to other hospitals within or outside the states.

5. Conclusion

This research explored the application of two fuzzy Delphi techniques, the MFD and TFD method to quantify medical expert opinions for ED metrics identification in a surge scenario. Comparing the results of the two fuzzy Delphi methods, twenty performance metrics were retained in the MFD method, while twenty-one was retained in the TFD method. The MFD method showed clear rankings with no ties because of the similarity and distance measure to obtain a consistency index of individual panel member opinion compared to the TFD method that applies only the distance measure. Finally, the three critical metrics ranked by the MFD method include ED beds, nurse staffing per patient seen, and employee satisfaction while the TFD method ranks employee fatigue, ED beds, and patients hospitalized as essential metrics for ED performance measurement. Also, some metrics were removed from the study, and the main reason is that most of the metrics did not achieve consensus as experts deemed them unimportant in measuring the ED's performance during a surge event. However, the research is still limited to biases from experts' opinions. The results in this paper are feasible and practical and can be used as a reference by hospital administrators for continuous quality improvement and measuring the performance of the ED during a disaster.

Acknowledgments

Funding: This work was supported by Blue Cross Blue Shield of Michigan Foundation (Grant #: 002934.PIRAP, 2020). The funding agency had no role in the study design, analysis, or decision to publish.

Author Declarations: All the relevant ethical guidelines have been followed and the WSU institutional review board (IRB) approvals have been obtained. The IRB #: IRB-19-11-1418

References

Chang, H., Cheng, M.A., et al. 2004. 'Hospital Ownership and operating efficiency: Evidence from Taiwan', *E. j. of oper. res.*, 159: 513-27.

Chang, Pao-Long, Chiung-Wen Hsu, and Po-Chien Chang. 2011. 'Fuzzy Delphi method for evaluating hydrogen production technologies', *International Journal of Hydrogen Energy*, 36: 14172-79.

Crist, Carolyn. 2020. 'Arizona, Florida, Texas Face Hospital Capacity', WebMD, Accessed July 18. https://www.webmd.com/lung/news/20200713/arizona-florida-texas-face-hospital-capacity.

Etu, Egbe-Etu Emmauel, Leslie Monplaisir, Celestine Aguwa, Suzan Arslanturk, Sara Masoud, Ihor Markveych, and Joseph Miller. 2020. "Identifying Indicators Influencing Emergency Department Performance during a Medical Surge via Modified Fuzzy Delphi Method." In *Working Paper*. Wayne State University.

Grasselli, Giacomo, Antonio Pesenti, and Maurizio Cecconi. 2020. 'Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: early experience and forecast during an emergency response', *Jama*.

Ho, Yu-Feng, and Hsiao-Lin Wang. 2008. "Applying fuzzy Delphi method to select the variables of a sustainable urban system dynamics model." In *Proceedings of the 26th International Conference of System. http://www.systemdynamics.org*. Citeseer.

Hsu, Yu-Lung, Cheng-Haw Lee, and Victor B Kreng. 2010. 'The application of Fuzzy Delphi Method and Fuzzy AHP in lubricant regenerative technology selection', *Expert Systems with Applications*, 37: 419-25.

Khalifa, Mohamed, and Ibrahim Zabani. 2016. "Developing emergency room key performance indicators: what to measure and why should we measure it?" In *ICIMTH*, 179-82.

Kuo, Ying-Feng, and Pang-Cheng Chen. 2008. 'Constructing performance appraisal indicators for mobility of the service industries using Fuzzy Delphi Method', *Expert Systems with Applications*, 35: 1930-39.

Madsen, Michael Moesmann, Andreas Halgreen Eiset, Julie Mackenhauer, Annette Odby, Christian Fynbo Christiansen, Lisa Kurland, and Hans Kirkegaard. 2016. 'Selection of quality indicators for hospital-based emergency care in Denmark, informed by a modified-Delphi process', *Scandinavian journal of trauma, resuscitation and emergency medicine*, 24: 11.

Murray, Thomas J, Leo L Pipino, and John P van Gigch. 1985. 'A pilot study of fuzzy set modification of Delphi', *Human Systems Management*, 5: 76-80.

Núñez, Alicia, Liliana Neriz, Ricardo Mateo, Francisco Ramis, and Arkalgud Ramaprasad. 2018. 'Emergency departments key performance indicators: A unified framework and its practice', *The International journal of health planning and management*, 33: 915-33.

Rezaei, Satar, Hamed Zandian, Akram Baniasadi, Telma Zahirian Moghadam, Somayeh Delavari, and Sajad Delavari. 2016. 'Measuring the efficiency of a hospital based on the econometric Stochastic Frontier Analysis (SFA) method', *Electronic physician*, 8: 2025.

Rezapour, Aziz, Zeynab Foroughi, Niusha Shahidi Sadeghi, Mehran Faraji, Alireza Mazdaki, Ali Sarabi Asiabar, Noureddin Niknam, Keivan Rahmani, and Saeed Mohammad-Pour. 2019. 'Identification of the most appropriate variables for measuring the efficiency of Iranian public hospitals: Using Delphi technique', *Journal of education and health promotion*, 8.

Rosko, Michael D, and Ryan L Mutter. 2008. 'Stochastic frontier analysis of hospital inefficiency: a review of empirical issues and an assessment of robustness', *Medical Care Research and Review*, 65: 131-66.

Saffie, N Amira M, and Khairul A Rasmani. 2016. "Fuzzy delphi method: Issues and challenges." In *2016 International Conference on Logistics, Informatics and Service Sciences (LISS)*, 1-7. IEEE.

Walters, Edgar, Shannon Najmabadi, and Emma Platoff. 2020. 'Texas hospitals are running out of drugs, beds, ventilators and even staff', *The Texas Tribune*.

Zarei, Leila, Farzad Peiravian, Mir Saman Pishvaee, and Bahar Aghababaei. 2019. 'The Scarce Drugs Allocation Indicators in Iran: A Fuzzy Delphi Method Based Consensus', *Iranian Journal of Pharmaceutical Research: IJPR*, 18: 1126.

Biographies

Egbe-Etu Etu is a Ph.D. candidate in the Department of Industrial & Systems Engineering at Wayne State University (WSU). His current research focuses on healthcare operations management with particular emphasis on improving the emergency department's performance during a surge event. He is an active member of Industrial Engineering and Operations Management (IEOM), Institute of Industrial & Systems Engineering (IISE), and SAVE International.

Leslie Monplaisir is a Professor in the Department of Industrial and Systems Engineering at WSU. He is the Lead Researcher and Director of the Product Development and Systems Engineering Consortium (PDSEC) at WSU. His research interests include Lean Product Development, Design for lean Systems and Services and Design reuse, New Product Technology Decision modeling, Product Architecture Optimization, and Healthcare Technology System Design. He has authored over 100 publications in these areas with funded research from NSF, Veterans Administration, and Ford Motor Company. Dr. Monplaisir joined the College of Engineering at WSU in the Department of Industrial and Manufacturing Engineering in 1996 from Florida A & M University, where he was a visiting assistant professor. He earned his Ph.D. in Engineering Management from the Missouri University of Science and Technology (MUST), a Master's in Computer Integrated Manufacturing from the University of Birmingham in Great Britain.

Celestine Aguwa is currently at WSU as an Associate Professor Research involved with research and teaching graduate courses in Industrial and Systems Engineering. His background includes lean and value methodology in product development and advanced manufacturing, customer voice analysis, and decision analysis modeling. He is currently working on several funded research projects in data analytics. He has a cross-functional industrial experience at Ford Motor Company and extensive professional experience as an Architect. Dr. Aguwa has a Ph.D. and MSIE in Industrial and Manufacturing Engineering from the University of Pittsburgh and Massachusetts, Amherst, respectively. He also has a B.Arch. Degree in Architecture from the University of Nigeria. Dr. Aguwa has several awards, including a patent, and has written several published papers. He is a member of the IISE, SAVE International, and Institute of Operation Research and Management Sciences (INFORMS).

Sara Masoud is an Assistant Professor of Industrial and Systems Engineering at WSU. Her research focuses on dynamic data-driven application systems by utilizing machine learning, simulation, and optimization models in different applications such as agro-industry, transportation, health care, and manufacturing. She is a member of IISE and INFORMS and has received the IISE Annual Meeting Best Paper Award in Data Analytics and Information Systems Track.

Suzan Arslanturk is an Assistant Professor in the Department of Computer Science at WSU. She was formerly an assistant professor at Ozyegin University in Istanbul, Turkey. She also worked as a research scientist for a healthcare

consulting company in Michigan while completing her Ph.D. at Oakland University in 2015. From 2011 to 2014, she was a research scholar at the OU Biomedical Engineering Lab. Her research interests are in the area of data mining and predictive modeling with applications in healthcare.

Ihor Markevych is a Master of Computational Data Science candidate in Language Technologies Institute, School of Computer Science at Carnegie Mellon University. His research interest is on differential modeling of semi-linear parabolic systems with composition method and convergence of iterations in the Trotter-Daletskii formula for nonlinear perturbation. He is an active member of the Association of Computing Machinery society.

Joseph Miller is an Emergency Medicine Physician at Henry Ford Hospital. He specializes in Emergency Medicine, Internal Medicine, and Clinical Research. Dr. Miller's research interest is on emergency neurological conditions, such as acute stroke, epilepsy, and traumatic brain injury. He also does research in hypertensive emergencies and teaches research methodology within the health system.