Introduction:

Sepsis is among the most common causes of hospitalization and death globally. It primarily affects patients with previous cardiovascular diseases like systolic heart failure. Patients with systolic heart failure are characterized by impaired cardiac output, reduced tissue perfusion, and increased susceptibility to infection, all factors that contribute to the high incidence of sepsis related complications and death. Understanding outcomes following sepsis in this subgroup is crucial for developing and implementing early detection and targeted treatment strategies.

Previous studies, such as that by Abou Dagher et al. (2018), investigated the outcomes of sepsis in systolic heart failure patients who were admitted to a tertiary hospital in Lebanon. They reported higher in hospital and 28 day mortality in patients with both conditions than in those without heart failure. Since sepsis remains a time-consuming and resource-intensive condition, reproducibility in studies helps evaluate their reliability and confirms whether the key outcomes identified persist when the same dataset is reanalyzed. This project aims to reproduce selected tables from the original study using data from the Dryad Digital Repository, and in the process, test the reproducibility and explore patterns that may inform future research on cardiovascular and infectious comorbidities. We chose to focus on reproducing results from Table 1 which focused on demographics and Table 5 which focused on the association of heart failure and hospital mortality of the patients being observed in the original study. Reproducing results of studies like this one further ensures reliability of the original study and can contribute to evaluation of study design and original methods used.

Methods:

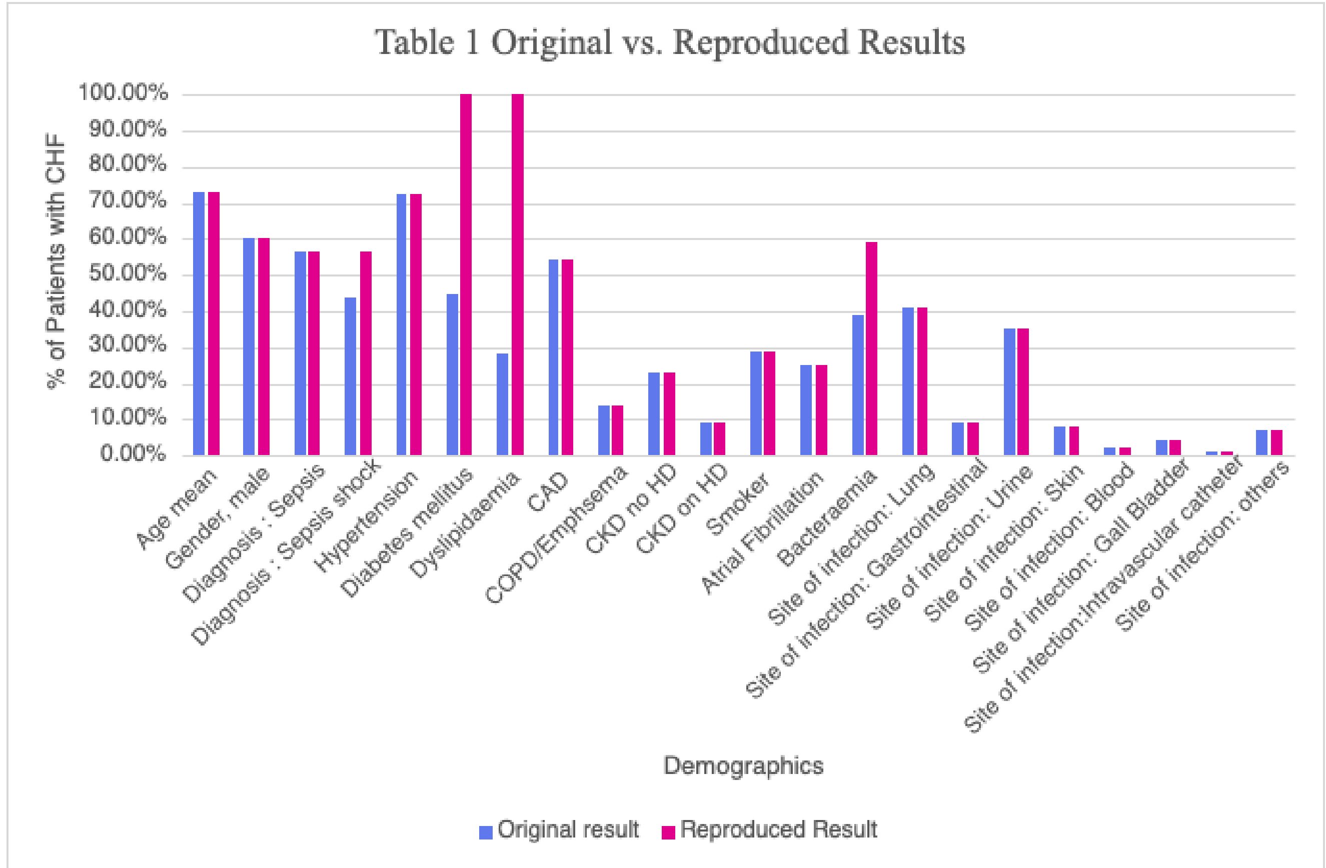
We used the dataset provided by the Dryad Digital Repository, which was linked directly to the original publication by Abou Dagher et al. (2018). This dataset contained anonymous patient information from a hospital in Lebanon. It includes comorbidities, demographic variables, diagnostic classifications, and mortality outcomes for individuals admitted to the emergency department with sepsis. The dataset was well organized, with each variable corresponding to a clinical or demographic characteristic, allowing us to reproduce the statistical results presented in the original study.

Before we started our analysis of this data we reviewed it to ensure completeness and consistency. Missing or ambiguous values were identified, and categorical variables were standardized to match those described in the original publication. Continuous variables such as age were summarized using measures of central tendency (mean and median) and dispersion (range and standard deviation), while categorical variables such as gender, comorbidities, and infection site were summarized using frequencies and percentages.

No additional data cleaning beyond standard formatting was required, as the dataset was already organized and ready for analysis.

Results:

Table 1 original results and reproduced results Bar graph comparison :



The comparison of the original results in Table 1 with the reproduced results shows that the reproduced data closely match the original study's findings. Most variables, such as age, gender, hypertension, diabetes, and sites of infection all showed similar percentages, showing accuracy within the reproduced results. Both datasets highlight that most patients were older and had several chronic conditions, especially hypertension and diabetes, with lung and urinary infections being the most common. The small differences that appeared were minimal and did not change the overall trend, confirming that the reproduced results accurately reflected the original result.

Table 5 original results and reproduced results Bar graph comparison :

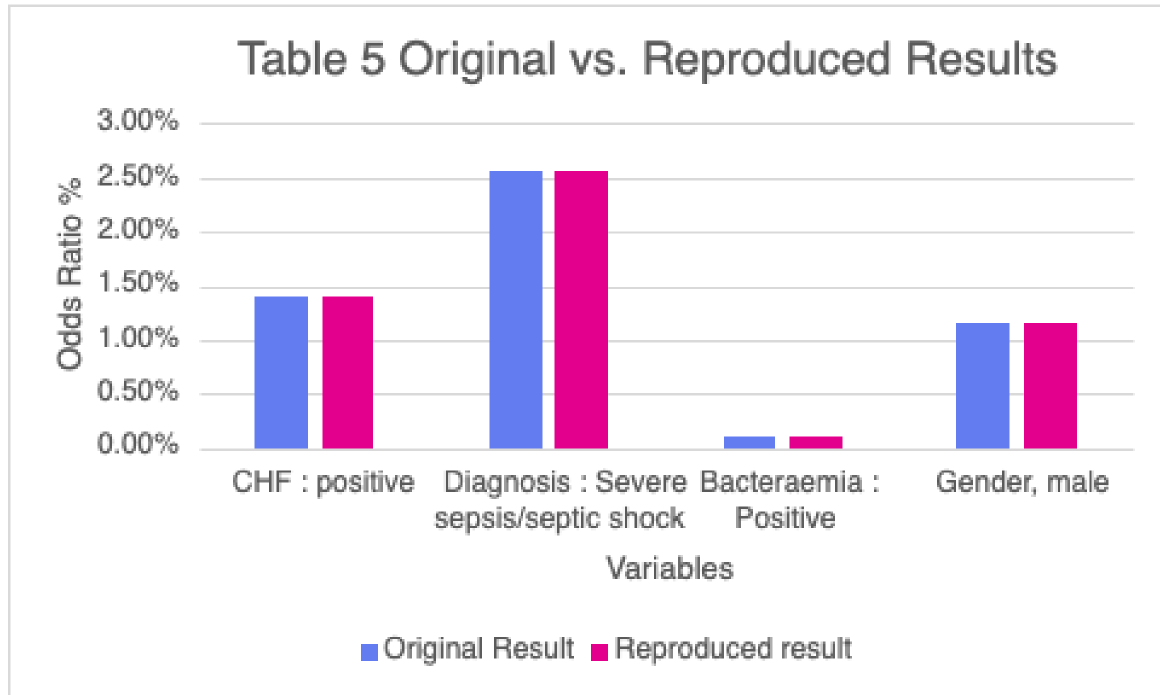


Table 5 shown above shows the comparison between the findings of the original results and the reproduced results, specifically examining the relationship between heart failure and hospital mortality. This table is based on a multivariable logistic regression model that assessed various clinical factors, including congestive heart failure (CHF), severe sepsis/septic shock, bacteraemia, and gender, to determine how each factor contributed to patient mortality.

The original results showed that patients who were diagnosed with congestive heart failure had a high hospital mortality. Patients with severe sepsis/septic shock had an even higher hospital mortality. While patients with bacteraemia actually had a lower risk. And patients that were male showed a slightly higher risk than patients with bacteraemia but not as of significant risk as Congestive Heart Failure or severe sepsis/septic shock.

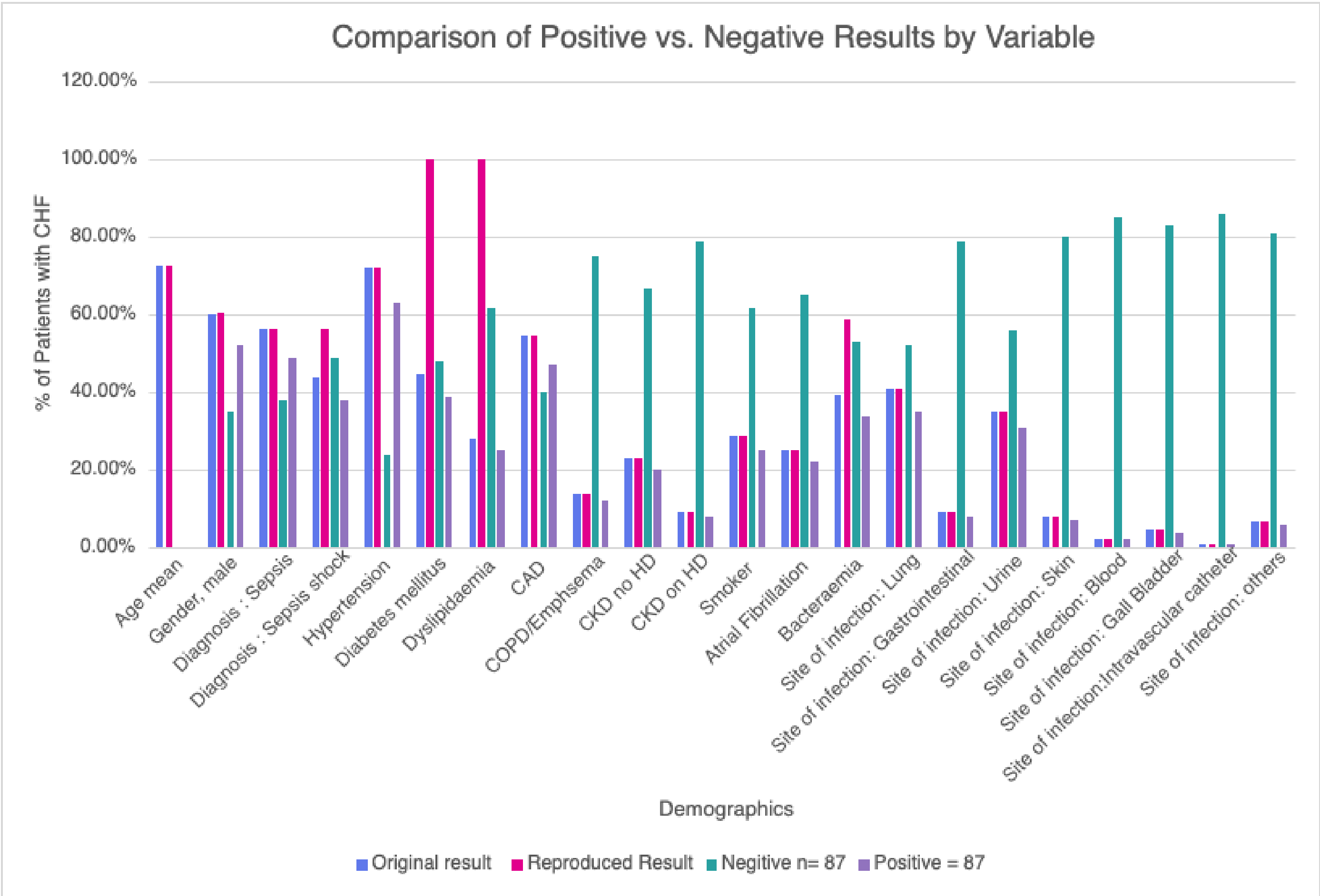
In the reproduced results, the percentages closely mirrored those of the original results. Patients with congestive heart failure had a hospital mortality rate of 1.408%, while patients with

severe sepsis/septic shock had a rate of 2.557%. The mortality rate for patients with bacteraemia was 0.1167%, and male patients had a rate of 1.155%. This comparison suggests that our reproduced results accurately reflect the original results.

Table 1:

Variable	CHF (total= 174)	CHF (Negative=87) (87 - positive)	CHF (Positive=87) ((%/100) x 87)
Age mean	72.80%	N/A	N/A
Gender, male	60.34%	35%	52%
Diagnosis : Sepsis	56.30%	38%	49%
Diagnosis : Sepsis shock	56.30%	49%	38%
Hypertension	72.40%	24%	63%
Diabetes mellitus	100%	48%	39%
Dyslipidaemia	100%	62%	25%
CAD	54.60%	40%	47%
COPD/Emphysema	13.80%	75%	12%
CKD no HD	23%	67%	20%
CKD on HD	9.20%	79%	8%
Smoker	28.70%	62%	25%
Atrial Fibrillation	25.30%	65%	22%
Bacteraemia	59%	53%	34%
Site of infection: Lung	40.80%	52%	35%
Site of infection: Gastrointestinal	9.20%	79%	8%
Site of infection: Urine	35%	56%	31%
Site of infection: Skin	8%	80%	7%
Site of infection: Blood	2.30%	85%	2%
Site of infection: Gall Bladder	4.60%	83%	4%
Site of	1%	86%	1%

infection:Intravascular catheter			
Site of infection: others	6.90%	81%	6%



We compared patients in the positive (n = 87) and negative (n = 87) groups to observe the various demographic factors, comorbidities, and infection sites. There were a total of 174 patients. The mean age was 72.8 years in both groups. Men made up just over half (60%) of the patients altogether.

Hypertension, coronary artery disease, and diabetes mellitus were the conditions that were seen the most. Hypertension was the most prevalent across these groups. 63% of positive

patients had hypertension and 24% of negative patients had it. 47% of the positive group had coronary artery disease (CAD) and 40% of the negative group had it. 39% of positive patients had diabetes mellitus and 48% of negative patients had it. Lastly, 25% of positive patients had dyslipidemia and 62% of negative patients had it.

Chronic kidney disease (CKD) with no hemodialysis was present in 20% of the positive patients and 67% of the negative patients. CKD with hemodialysis was less common in both the positive and the negative groups, 8% and 79% respectively. Chronic Obstructive Pulmonary Disease (COPD) or emphysema was present in 12% of positive and 75% of negative patients. A quarter (25%) of the positive patients were former or current smokers and 62% of the negative patients were current or former smokers. Arterial fibrillation was present in 22% of positive patients and 65% of negative patients.

Sepsis was identified in almost half (49%) of the positive patients and in 38% of negative patients. Septic shock turned out to be more common in the negative group with 49% and only 38% in the positive patients. Bacteremia was found in 34% of positive patients and 53% of the negative patients.

The most common site of infection was the lungs in 35% of the positive group and 52% of the negative group. After the lungs were urinary tract infections (31% for positive and 56% for negative) and gastrointestinal infections (8% for positive and 79% for negative). Similarly, skin infections were common in the negative group as well (7% for positive and 80% for negative). Other infection sites observed were blood, gallbladder, intravascular catheter, and miscellaneous. The prevalence in each of the previously listed sites was significantly higher for the negative groups compared to the positive groups. These findings are represented in the graph above.

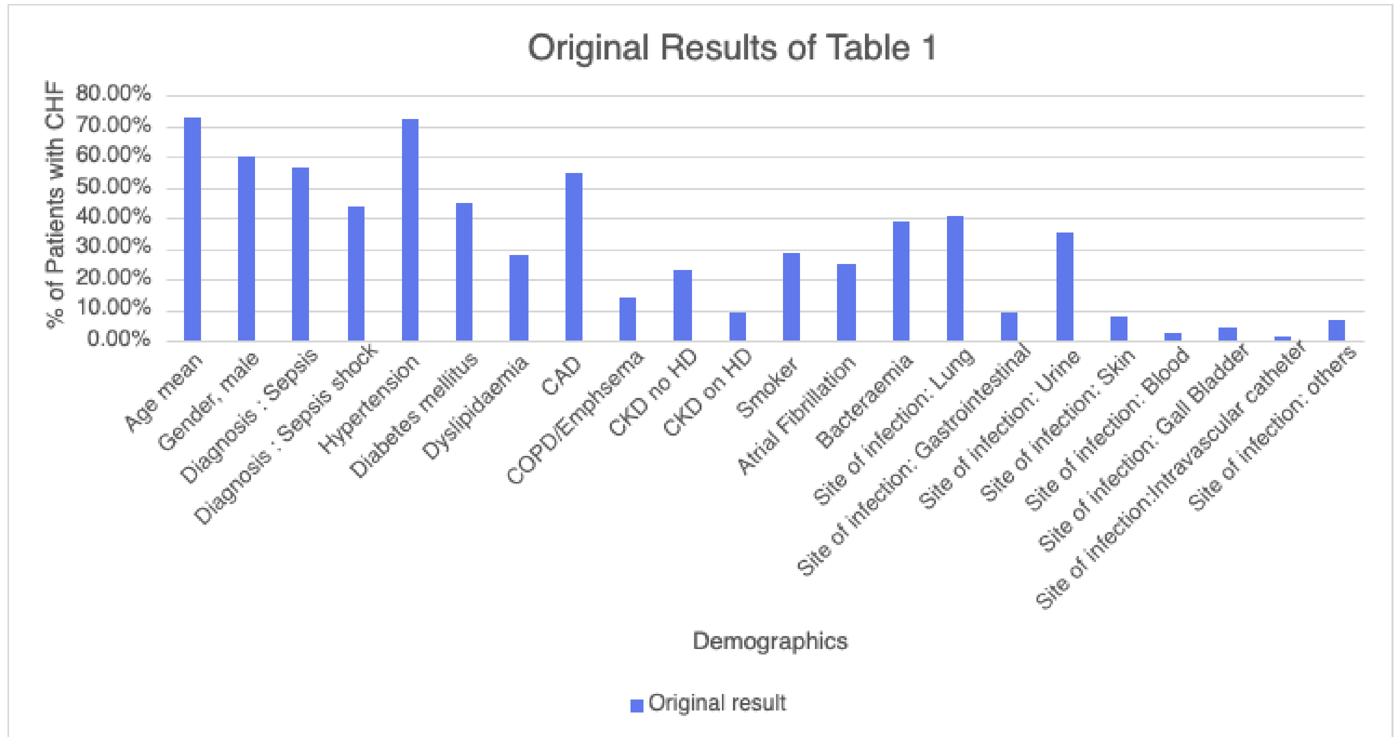
While most of the reproduced results were closely approximated to the original data, there were some discrepancies between the two data sets. Variations could be due to rounding differences or incomplete/incorrect information. If the original paper/source gave exact calculation details, reproduced results may be more accurate.

Table 5:

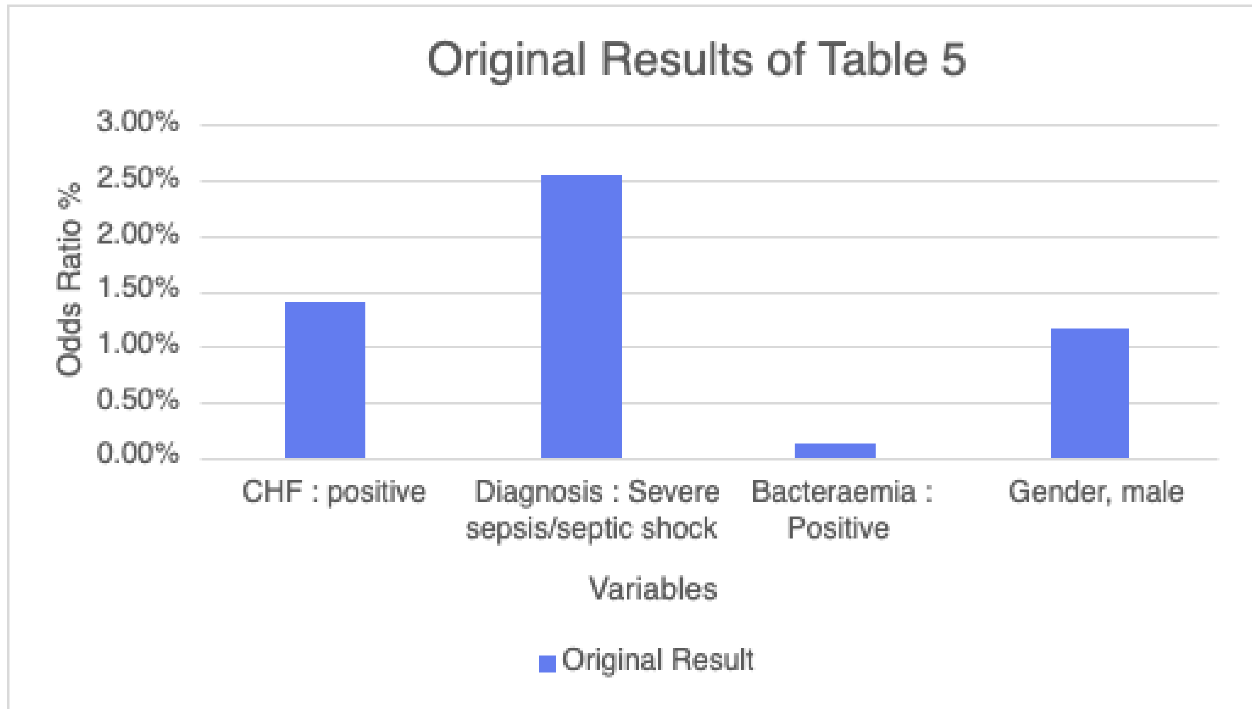
Variable	Hospital Morality	Odds Ratio	P-Vaule
CHF	High mortality	2.45%	0.01
Severe Sepsis/ Septic Shock	High mortality	4.45%	<0.001
Bactereamia	Lower mortality	0.29%	0.001
Male Gender	Slight increase	2.01%	0.06

The measure of association table shown above (Table 5) demonstrates the variables associated with heart failure and hospital mortality. The results presented in the table indicate that both congestive heart failure and severe sepsis/septic shock are associated with a higher risk of hospital mortality, as evidenced by a low P-value of 0.01, which demonstrates statistical significance. In contrast, bacteremia was associated with a lower rate of hospital mortality, resulting in a significantly lower P-value. Additionally, the data showed that male gender was associated with a slight increase in mortality, with a P-value of 0.06, suggesting that this finding is not statistically significant.

Figures of original data within the paper :



The Original Results of Table 1 describe the characteristics of the patients in the study. Most participants are older adults, with an average age of around 73 years, and a higher ratio of men than women. Many patients had chronic illnesses such as hypertension, diabetes, coronary artery disease, and kidney disease. Among those with heart failure, respiratory and urinary infections were the most common, while gastrointestinal and skin infections were less frequent. Table 1 indicates that patients with heart failure experience more health complications and exhibit slightly different patterns of infection compared to those without heart failure.



The Original Results of Table 5 show the findings from a multivariable analysis examining the relationship between heart failure and hospital mortality. When controlling for different variables like age, gender, and other health conditions, the results showed that patients with congestive heart failure had about 2.5 times higher rates of dying in the hospital. Patients diagnosed with severe sepsis or septic shock had the highest risk, with about 4.5 times greater of a chance of experiencing hospital mortality. In contrast, patients with bacteremia had low odds of dying in the hospital, therefore these patients had a lower risk of experiencing hospital mortality. Overall, Table 5 shows that heart failure and severe infection were the main predictors of hospital death, even after accounting for other health factors.

Discussion:

Our reproduced results closely support those of the original source, Abou Dagher et al. (2018). These results confirm that patients with systolic heart failure who present with sepsis experience significantly higher mortality rates and more frequent multi-organ involvement than those without heart failure. The consistency between our reproduced data and the original publication increases the reliability of the original findings. Given the small differences in percentages-which might have resulted from rounding, data format, or inclusion exclusion factors the direction and magnitude of the findings remained comparable. This implies that the dataset on the Dryad repository was well maintained and gives a reliable basis for secondary analysis and validation studies.

The reproducibility of this paper was generally strong, with a well-documented dataset containing variable names and accessible metadata. However, several barriers limited the complete replication of all statistical outputs. First, there was a lack of exact definitions for some diagnostic categories, such as what exactly distinguished "severe sepsis" from "septic shock," making variable coding ambiguous. Notably, a very detailed codebook would have been welcomed to fully reproduce the parameters of logistic regression and estimates of hazard ratios in the original study. A further limitation was that the original R script and analytic code were not available, necessitating reasonable assumptions on our part about how various calculations, statistical models, and transformations of variables had been done. This points to some limitations in the comprehensiveness of the documentation and transparency of published research.

In order for us to further enable the reproducibility of this data, the original authors could have included a few key components. First, the complete code used for data cleaning and analysis. Second, clearly defined diagnostic criteria and categorical groupings. And third,

explicit documentation of how missing data were handled. For example, when it came to Figure 1 gastrointestinal is a demographic that was missing data. It should have been noted why gastrointestinal was left out of the data. By having clear reasoning for data choices reproducibility can be improved. Greater transparency in this area is necessary to improve scientific integrity and advance evidence-based reasoning.

In summary, this study serves to underscore the importance of open data sharing and methodological transparency in epidemiological research. Reproducibility and validation ensure that conclusions from data are credible and relevant to clinical practice. Confirming the reproducibility of Abou Dagher et al. (2018) enhances certainty in the reported association between sepsis and poor outcomes among patients with systolic heart failure.

References:

Abou Dagher G, Hajjar K, Khoury C, El Hajj N, Kanso M, Makki M, Mailhac A, Bou Chebl R. Outcomes of patients with systolic heart failure presenting with sepsis to the emergency department of a tertiary hospital: a retrospective chart review study from Lebanon. BMJ Open. 2018 Aug 1;8(7):e022185. doi: 10.1136/bmjopen-2018-022185. PMID: 30068620; PMCID: PMC6074621.