A COMPUTATIONAL ANALYSIS OF DIFFERENCES IN THERAPY BETWEEN BENCHMARK AND NON-BENCHMARK HOSPITALS FOR PATIENTS WITH ACUTE MYOCARDIAL INFARCTIONS

Raphael Bahati[1], Michael Bauer[2] and Femida Gwadry-Sridhar[1]
[1]*I-THINK Research, Lawson, 801 Commissioners Rd., Suite B3041, London, ON, Canada*
[2]*Department of Computer Science, The University of Western Ontario, London, ON, Canada*

Keywords: Acute myocardial infarction, Predicting health outcomes, Mathematical modeling, Cluster analysis.

Abstract: Acute Myocardial Infarction (AMI) remains a leading cause of mortality in most industrialized nations. Mortality rates for AMI patients are often used as a measure of the overall effectiveness of care provided by hospitals. Age, gender, and severity adjusted, the mortality rates within Canada have been shown to vary significantly from province to province. Some studies, for example, have shown significant variations between counties, even when adjacent to each other. In this paper, we present an approach aimed at understanding the causes of this variability by investigating the extent to which evidence-based therapies and processes within hospitals might be affecting mortality rates. We use cluster analysis to identify beneficial therapies and processes responsible for the improvement in treatment outcomes (as measured in terms of standardized mortality ratio) in benchmark compared to non-benchmark hospitals.

1 INTRODUCTION

Acute Myocardial Infarction (AMI), commonly known as a heart attack, is caused by a sudden deprivation of blood circulation to parts of the heart mainly as a result of a blockage of coronary artery. The shortage of oxygen often causes permanent myocardial (heart) tissue damage or death. In 1996, 38,000 myocardial infarctions were reported in Canada, the majority due to coronary artery disease (HSFC, 2010). Approximately 15% of AMI sufferers died and 23% were readmitted within the first year post AMI (CHSSS, 2000). Consequences of coronary artery disease include morbidity such as angina, congestive heart failure and arrhythmias. These disease related morbidities often lead to significant disability and economic impact from diminished productivity and ongoing health care costs.

The mortality rates for AMI have been shown to vary significantly from province to province and within provinces from county to county. The Institute for Clinical Evaluative Sciences published a Cardiovascular Atlas, which compiled the rate of mortality from AMI by district or county in Ontario, Canada (Basinski et al., 1999). It showed significant variance between counties even when adjacent to each other. In southwestern Ontario, the rate of age and sex adjusted cardiovascular mortality per 100,000 population greater than age 20 ranged from 368.7 (Middlesex County) to 526.5 (Kent County). Another study by the Canadian Cardiovascular Outcomes Research Team (CCORT) published a nationwide perspective of AMI mortality rates from seven provinces, adjusting for age and sex. The adjusted mortality rates ranged from a low of 10.2% in Alberta to a high of 13% in Saskatchewan (Tu et al., 2003). What this demonstrates is that age or sex differences did not account for the variations in the rate of AMI mortality. Furthermore, the variability in mortality could be seen when examined at the level of individual hospitals. Severity adjusted, (expected) mortality ranged between 10 and 17% in a representative group of hospitals in Ontario, Alberta and Newfoundland, but observed mortality had nearly four times as much variation (between 3.5 and 12.5%). Thus factors such as age, co-morbid disease and severity of presentation did not entirely account for the differences in treatment outcomes.

This begs the question as to why we see significant variability in outcome indicators such as AMI in-hospital mortality and ICU mortality. AMI is divided into 2 major categories based on electrocar-

Bahati R., Bauer M. and Gwadry-Sridhar F..
A COMPUTATIONAL ANALYSIS OF DIFFERENCES IN THERAPY BETWEEN BENCHMARK AND NON-BENCHMARK HOSPITALS FOR PATIENTS WITH ACUTE MYOCARDIAL INFARCTIONS .
DOI: 10.5220/0003157502540259
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2011), pages 254-259
ISBN: 978-989-8425-34-8

diogram (ECG) diagnostic test for detecting heart-muscle damage: ST elevation and non-ST elevation myocardial infarction (MI). ST elevation MI is caused by an acute thrombotic occlusion of a major coronary vessel while Non-ST elevation MI is often associated with partial closure of an epicardial vessel or with diffuse coronary artery disease. Effective therapies exist for both categories of AMI. One possible explanation for the variability in treatment outcomes is the extent to which different kinds of therapies available at each hospital might be affecting mortality rates. Specific treatment issues may be reflected in the percentage of patients who receive medications or interventions known to improve the chance of survival.

In this paper, we investigate the impact of the different kinds of therapies on the mortality rates by identifying beneficial therapies and processes responsible for improving treatment outcomes in benchmark versus non-benchmark hospitals. Our focus is on ST elevation MI and the specific evidence-based therapies to achieve rapid reperfusion of the occluded vessel. The rest of this paper is organized as follows. We begin in Section 2 with an overview of the data source used in the analysis, describing patient composition within ICUs involved as well as the benchmark methodology used to distinguish benchmark from non-benchmark hospitals. We then describe our analytical approach for identifying beneficial therapies and processes responsible for improving treatment outcomes in benchmark versus non-benchmark hospitals in Section 3. We conclude with a summary of the implications of our study and describe possible directions for future work in Section 4.

2 DATA SOURCE

The Critical Care Research Network is a network of ICUs within Ontario established to conduct evidence-based research within both teaching and community hospitals and to facilitate research transfer to the decision-makers within these settings. The Network has been collecting a Minimum Data Set (MDS) since January 1995. The MDS currently contains over 125,000 records from 45 hospitals from across Canada. The dataset contains hospital and ICU admission and discharge dates, hospital outcome, ICU admitting diagnosis, and physiologic data for calculating an illness severity score on the day of ICU admission. Every admission to the ICU is recorded. Acute myocardial infarction is one of the specific diagnoses captured in the dataset. Sites collect data on all ICU admissions with $> 90\%$ of records containing complete data. Strengths of the database include the APACHE (Acute Physiology And Chronic Health Evaluation) II score, collected as part of the MDS, which has been validated as an index of severity and can be used to adjust for illness severity when comparing outcomes between coronary care units. This is the most widely used method worldwide for risk adjustment of ICU patients. Also, the diagnosis has to be determined during the first 24 hours of ICU admission and the patient location prior to ICU admission is recorded. Thus, patients with AMI as a secondary diagnosis (e.g. post-operative) can be excluded.

2.1 Site and Patient Selection

Sites were included in the analysis if they were a community hospital (since most teaching hospitals have separate coronary care units) and at least 10 cases per year were recorded in the database. Although only ICUs were included in this study, this represented the majority of community practice, since only 8 of 28 Critical Care Research Network (CCR-Net) community hospitals reported a coronary care unit separate from the main intensive care unit in our most recent survey. Patients were included in the analysis if they had a diagnosis of acute myocardial infarction, were admitted directly from the emergency department to the ICU, and were at least 16 years old. The composition of ST elevated MI patients within hospitals and the corresponding demographics are summarized in Table 1.

2.2 Benchmark Methodology

Objective methodology to identify best practice has been described in (Weissman et al., 1999) and used in randomized controlled trials for quality improvement (Kiefe et al., 2001). This methodology was implemented using risk-adjusted mortality. Thus, the predicted risk of death is calculated for each patient using the APACHE II risk prediction model (Knaus et al., 1985). The average predicted risk of death is then determined for each ICU and compared to the actual mortality rate as a ratio (SMR, standardized mortality ratio), with an adjustment for small sample sizes by adding 1 to the numerator and denominator. Sites were then ranked in order of the SMR. Starting with the highest ranked site, sites were added to the benchmark group until at least 10% of the total patient pool was included. A pooled SMR was generated for the overall benchmark group of patients. Confidence intervals were then generated according to the method of Hosmer and Lemeshaw (Hosmer and Lemeshow, 1989) and used to group ICUs into benchmark versus non-benchmark hospitals.

Table 1: Demographics of ST elevated MI patients. HLOS = Hospital Length of Stay, Values: mean (standard deviation).

	Benchmark Hospitals			**Non Benchmark Hospitals**			
Hospital ID	1629	3411	4067	1754	1768	1853	3587
Patients	25	25	22	25	25	25	25
Gender[F/M]	5/20	8/17	4/18	4/21	9/16	10/15	6/19
Age	54.52 (10.32)	60.56 (14.02)	55.05 (10.96)	57.32 (11.17)	67.24 (10.23)	60.20 (10.21)	59.48 (10.53)
HLOS	4.08 (1.66)	5.16 (2.87)	7.36 (4.47)	5.04 (3.92)	7.28 (3.66)	4.64 (2.78)	3.88 (1.86)
Drugs	6.36 (1.08)	5.88 (1.79)	5.91 (1.27)	5.68 (0.99)	5.76 (1.51)	6.60 (1.53)	5.68 (1.49)
APACHE II	7.91 (1.66)	9.88 (3.27)	8.62 (4.73)	8.92 (3.26)	8.56 (2.68)	8.28 (3.06)	7.88 (3.38)

3 ANALYSIS

In this section, we describe the approach taken in understanding the variability in health outcomes of patients with AMI as a result of being treated at either benchmark or non-benchmark hospitals. In particular, we examine how Evidence Based Medicine (EBM) therapies and processes affected treatment outcomes as measured in terms of the standardized mortality ratio. The aim is to identify beneficial therapies by comparing treatment outcomes in benchmark versus non-benchmark hospitals.

3.1 Approach

Our analytical approach involved two key steps. We used the analysis of variance (ANOVA) to test the effects of the different factors (i.e., therapies and tests) on the outcome measure (i.e., standardized mortality ratio). We used ANOVA as a starting point for identifying therapies of interest (Section 3.2), whose results were then fed onto a k-means clustering algorithm (Kiefe et al., 2001) for partitioning patients data (Section 3.3). Thus, instead of building clusters using all therapies and tests in Tables 2, 3, and 4 (25, in this case), we only included those that significantly distinguished benchmark from non-benchmark hospitals. This allowed us to eliminate irrelevant therapies and tests resulting in greater cluster stability.

The use of clustering in our analysis served two purposes. First, it allowed us to verify and test the accuracy of the identified therapies in classifying patients as having been treated at either a benchmark or a non-benchmark hospital. Second, it provided a natural way of partitioning patients into groups based on treatment characteristics. Consequently, we were then able to compare characteristics of individual clusters to determine which therapies and processes were beneficial and which were not.

3.2 Analysis of Variance

A multi-factor analysis of variance (ANOVA) was used to test the effects of different therapies and tests on the treatment outcome as determined by whether a patient was treated at a benchmark versus a non-benchmark hospital (see Section 2.2). Each factor (therapy) consisted of two levels (denoted by 1 and 0) indicating whether or not a particular treatment was administered to a patient during a hospital visit. From this analysis, a number of therapies and tests emerged as significant (at $p < 0.05$) and are highlighted in Tables 2, 3, and 4. They include one pharmacologic therapy (A), four non-pharmacologic therapies (B, C, D, and E), and five diagnostic tests (F, G, H, I, and J).

Table 2: Pharmacologic Therapies.

	Therapy	**P**
	ASA	0.36
	Reperfusion	0.38
	Thrombolytics	0.14
	Anti-thrombotic	0.17
	Statin	0.41
	Beta blocker	0.98
A	ACE-Inhibitor	0.04
	Statin preprint	0.73

3.3 Cluster Analysis

Having identified the therapies of interest, we then used cluster analysis to group patients such that each group (or cluster) consisted of patients that underwent a combination of treatments with similar characteristics. In particular, a k-means clustering algorithm was used to partition patients based on which therapies, from among the therapies identified as significant in the previous section, they received while in hospital. Thus, given a set of patients (see, for example, Table 5), k-means clustering partitioned the patients into k groups such that each patient belonged to

Table 3: Non Pharmacologic Therapies.

	Therapy	P
	Patients monitored for 48-72 hours	0.40
	Patients reassessed at 48-72 hours to remove or continue monitoring	0.34
	Patients reassessed with trained personnel observing	0.47
	Protocol for early mobilization	-
	Staged or formal mobilization of patient	-
	Pharmacist input documented	0.43
B	Pharmacist participation in rounds	<0.01
C	Disease education documented	<0.01
D	Education material made available for patients and families	0.03
E	Discharge planning documented	0.04
	Discharge formalized with planning	0.86

Table 4: Key Diagnostic Tests.

	Diagnostic Test	P
	Chest pain unit or rapid triage for chest pain	0.38
F	Assessment of infarction size	0.05
G	Assessment of residual ischemia pre-discharge	<0.01
H	Referral made for cardiac catheterization	<0.01
I	Cardiac Care Network forms in chart	<0.01
J	Formal guidelines for decision making regarding cardiac catheterization	<0.01

a cluster with similar treatments characteristics. Each patient's information, in this case, consisted of a set of binary numbers each corresponding to a therapy in {A,B,...,J} indicating whether (1) or not (0) a particular therapy was administered.

Table 5: Sample ST elevated MI patients data.

	Treatments and Tests				
Patient	**A**	**B**	...	**I**	**J**
1	1	0	...	0	0
2	1	1	...	1	1
			. . .		
n	0	0	...	0	1

The k value for the algorithm, which determines the number of groups (or clusters) to be created, was chosen as follows. We experimented with several cluster configurations (where $2 \leq k \leq 20$) such that k resulted in the most accurate classification of patients. Accuracy, in this case, described the patient composition of each cluster based on whether they were treated at a benchmark or a non-benchmark hospital. Thus, an accuracy measure of, say, 0.8 meant that, for every cluster formed, at least 80% of its patients were treated at the same kind of hospital. As such, the aim was to group patients into the smallest possible number of clusters while ensuring the most accurate grouping of patients based on whether they were treated a benchmark versus a non-benchmark hospital. In this approach, five clusters were formed as shown in Table 6 with cluster 2 being the least accurate (at 73.91%) and clusters 4 and 5 being the most accurate (at 100%).

Table 6: Clusters Prediction.

	% Patients Composition	
Cluster	**Benchmark**	**Non-Benchmark**
1	75.51	24.49
2	26.09	73.91
3	92	8
4	0	100
5	0	100

Of particular interest, clusters 4 and 5 (highlighted in Table 6) had a 100% accuracy in classifying patients as having been treated at non-benchmark hospitals. Looking more closely at the patient composition within the clusters reveals some interesting characteristics as Table 7 shows. For example, of the 25 patients in cluster 3, 92% were treated at Hospital 1629 while 8% were treated at Hospital 1768. Note also that, all the patients in cluster 4 were treated at Hospital 3587, while all the patients in Hospital 1754 ended up in cluster 5. This suggests, to some extent, a correlation between therapies and hospitals.

Table 8 shows the percentage of patients that un-

Table 7: Percentage patient composition within clusters.

	Benchmark Hospitals			Non Benchmark Hospitals				
Cluster	**1629**	**3411**	**4067**	**1754**	**1768**	**1853**	**3587**	# Patients
1	0	36.73	38.78	0	22.45	0	2.04	49
2	4.35	15.22	6.52	0	26.09	47.83	0	46
3	92	0	0	0	8	0	0	25
4	0	0	0	0	0	0	100	24
5	0	0	0	89.29	0	10.71	0	28
# Patients	25	25	22	25	25	25	25	

Table 8: Percentage of patients that underwent specific therapies and tests within clusters.

	Pharm.	Non-Pharmacologic				Diagnostic Tests					
Cluster	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**Accuracy**
1	63.27	2.04	46.94	100	89.8	93.88	55.1	44.9	4.08	100	0.76
3	48	0	60	36	52	92	76	100	100	8	0.92
2	56.52	0	8.70	19.57	6.52	80.43	17.39	93.48	78.26	95.65	0.74
4	29.17	100	29.17	20.83	91.67	91.67	95.83	95.83	70.83	100	1
5	67.86	89.29	0	0	3.57	100	39.29	57.14	39.29	0	1

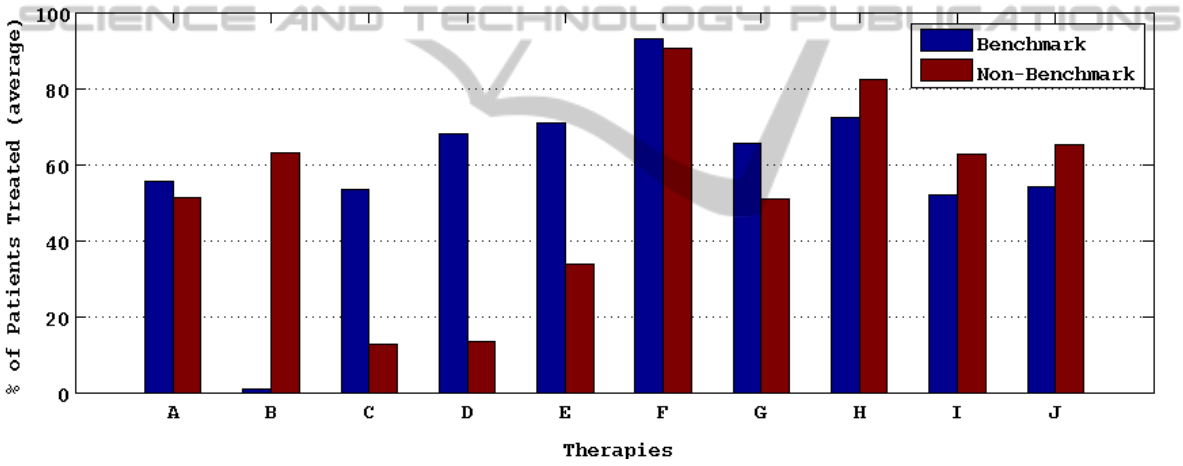

Figure 1: Comparing therapies and tests between benchmark and non-benchmark clusters.

derwent specific therapy treatments and tests within each cluster. The clusters are grouped into benchmark (clusters 1 and 3) versus non-benchmark (clusters 2, 4, and 5) based on the patient composition within individual clusters (see Table 6). For example, cluster 1 is considered a benchmark cluster since 75.51% of its patients were treated at benchmark hospitals whereas cluster 2 is considered a non-benchmark cluster. In order to compare individual therapies and tests between benchmark and non-benchmark clusters, we computed the average percentage of patients that underwent each specific treatment in {A,B,...,J} as shown in Figure 1.

Since the aim was to identify which therapies were beneficial in terms of the improvement in treatment outcomes (assuming, of course, that therapies and tests did not contribute to the worsening of treatment outcome as measured in terms of standardized mortality ratio), several conclusions can be drawn from the results. On the one hand, therapy B, which corresponded to pharmacists participation in rounds (see Table 3), does not appear to be beneficial as it was mostly administered to patients treated at non-benchmark hospitals, which performed worse based on standardized mortality rates. On the other hand, we see much larger differences in the percentage of patients that underwent therapies C, D, and E in favor of benchmark hospitals. This suggests that such therapies and tests may have contributed to the improvement in treatments outcomes.

A t-test comparison of the percentages of patients that underwent specific therapies between benchmark and non-benchmark clusters at a 95% confidence shows that non-pharmacologic therapy C was the only significant predictor of the variations in health outcomes ($p = 0.03$). Worth pointing out, however, is the fact that not all benchmark hospitals (or non-benchmark hospitals for that matter) offered the same kinds of therapies. As such, it could be a combination of different treatments that may have been responsible for the overall improvement in treatment outcomes. Cluster analysis is only the first step in helping us identify some of these characteristics.

4 CONCLUSIONS

The variability in mortality rates of AMI patients between hospitals is due to the differences in the kinds of therapies and tests administered to patients at different hospital locations. In this paper, we have presented an approach for identifying beneficial therapies and processes responsible for improving treatment outcomes of ST elevation MI patients. To achieve this, we first used the analysis of variance (ANOVA) to test the effects of the different therapies on the outcome measures. This allowed us to identify therapies of interest, which we then used on a k-means clustering algorithm to group patients. In this approach, patients belonging to the same cluster underwent similar therapy treatments while in hospitals with similar outcomes. Consequently, we were able to compare treatment characteristics within clusters to determine which therapies were beneficial based on the differences in the percentage of patients treated at benchmark versus non-benchmark clusters.

Several conclusions could be drawn from the analysis presented in this paper. First, therapy B (pharmacists participation in rounds) does not appear to have any benefit in the improvement of treatment outcomes as it was mostly administered to patients treated at non-benchmark hospitals, which performed worse based on the standardized mortality ratio. Second, therapy C (disease education documentation) was the only significant predictor in the overall improvement of treatment outcomes in benchmark versus non-benchmark hospitals.

In our current approach, we have utilized the APACHE II risk prediction model to determine the standardized mortality ratio, which we computed by comparing the predicted risk of death as determined by the APACHE II score to the actual hospitals' mortality rates. Future work includes the incorporation of other outcome indicators such as hospital length of stay within the cluster analysis. An interesting question, which deserves further investigation, is whether or not patients who leave hospital earlier tend to have a worse clinical outcome compared to those who stay in hospital longer. For example, is it possible that hospital length of stay might indirectly be impacting the variations in mortality rates? This is the focus of our future investigation.

ACKNOWLEDGEMENTS

This research was funded by the Canadian Institute of Health Research (CIHR) and the Critical Care Research Network (CCR-Net).

REFERENCES

Basinski, A. S., Naylor, C. D., and Slaughter, P. (1999). Hospitalization for Cardiovascular Medical Diagnoses, Cardiovascular Health and Services In Ontario. In *Institute for Clinical Evaluative Services and Heart and Stroke Foundation (Ontario)*, Toronto, Canada. ICES Atlas.

CHSSS (2000). The Changing Face of Heart Disease and Stroke in Canada. In *Canadian Heart and Stroke Surveillance System*, Ottawa, ON, Canada. Health Canada.

Hosmer, D. W. and Lemeshow, S. (1989). *Applied Logistic Regression*. John Wiley & Sons, Inc., New York, NY, USA.

HSFC (2010). Heart & Stroke Foundation of Canada. http://www.hsf.ca/research/en/general/home.html.

Kiefe, C. I., Allison, J. J., and Williams, O. D. (2001). Improving quality improvement using achievable benchmarks for physician feedback: a randomized controlled trial. In *JAMA*, volume 285, pages 2871–2879, Ottawa, ON, Canada.

Knaus, W. A. et al. (1985). APACHE II: a severity of disease classification system. In *Critical Care Medicine*, volume 13, pages 818–829.

Tu, J. V. et al. (2003). Outcomes of Acute Myocardial Infarction in Canada. In *Canadian Journal of Cardiology*, volume 19, pages 893–901.

Weissman, N. W. et al. (1999). Achievable benchmarks of care: the ABCs of benchmarking. In *J Eval Clin Pract*, volume 5, pages 269–281. Health Canada.