Computational Modeling to Predict Demand for Chest Pain Management

Authors: Ruoming Jin PhD, Wei Xiong PhD, Brigitte Piniewski MD

Abstract— Each time an individual presents at the emergency room with chest pain, this is an expensive problem which uses multiple high cost hospital units such as the emergency room, radiology, the cath lab, cardiac intensive care unit and so on. Currently, most hospitals are poorly able to predict the demand for these services and thus may under staff or over staff these departments on any given day. Imagine if a computational model could deliver a "community chest pain index" and enable services to gear up or gear down in anticipation of a more predictable demand for their expensive services.

To solve this problem, PeaceHealth Laboratories connected live laboratory testing volumes with the computational expertise of Kent state University. A reasonable surrogate quantifying the level community chest pain management demand is troponin laboratory testing rates per date per location. The Kent State team used the first two years of troponin volumes as training data and the last half year as test data for three communities. The error measure was as low as 13.2% in one community. Males were consistently more predictable than females. The day shift variability was greater than night shift, and the week days were more predictable than weekends. A number of adaptive modeling and stratified modeling incorporating a sliding rule was used to further shrink the error measure which could also be broken down by the individual day.

Finally, non-traditional community influencers such as sport games, weather changes, unemployment trends, and crime rates may influence the "community chest pain index". Our work is currently exploring the additional predictive capacity afforded to the models when these non-traditional yet potentially high yield dynamic contributors are also taken into account.

In the future, the ability to predict a "high or low chest pain index" will enable evidence-based or "computational model-based" staffing for many expensive hospital areas thus significantly reducing costs. Also new theories of individual human behavior will depend upon a solid understanding of the contextual ebb and flow of dynamic background group behaviors.

This work will help lay the ground work for Experience-based medicine which will be fundamental to supporting experience-based theories of human health behaviors at large.

R. Jin is with the Department of Computer Science, Kent State University, Kent, OH 44242 USA (corresponding author to provide phone: 1-330-672-9063; e-mail: jin@cs.kent.edu).

W. Xiong is with the Department of Computer Science, Kent State University, Kent, OH 44242 USA (e-mail: wxiong@cs.kent.edu).

B. Piniewski is Chief Medical Officer with Peacehealth Laboratories USA (e-mail: bpiniewski@peacehealthlabs.org).

I. Introduction

The objective of this research is to design models that can accurately predict the expected daily demand for troponin testing in three communities. We will analyze the time series for each community separately, by using basic time series techniques, since the only information currently available is the number of troponin lab tests.

The rest of paper is organized as follows. In Section II we describe the dataset we used in our study. In Section III, we present auto-vector regression methods to predict the expected demand of daily troponin tests. In Section IV, we focus on adaptive methods. In Section V, we discuss the results. In Section VI, we conclude.

II. Dataset

The Troponin lab data dates from 01/01/2011 to 05/31/2013, with a total of 882 days. It was collected through the following six communities: Ketchikan, Alaska; Cottage Grove, Oregon; Eugene/Springfield, Oregon; Florence, Oregon; Longview, Washington; Vancouver, Washington & PDX; Whatcom, Washington.

We used SAS to extract the daily visits from the original data. The column "ORDERED" in the original Excel file is selected as the dates when the patients visit the hospital. The following is a table for the total test volume and average daily test volume for each of the regions in those 882 days:

Communities	Total Volume	Average Daily volume
Ketchikan (Alaska)	3,984	5
Cottage Grove (Oregon)	2,542	3
Eugene/Springfield (Oregon)	**35,447**	40
Florence (Oregon)	4,191	5
Longview (Washington)	**30,580**	35
Vancouver (Washington)	237	0
Whatcom (Washington)	**58,823**	67

As we can see from this table, three of the communities have relatively large testing volumes: Eugene (Oregon), Longview (Washington), and Whatcom (Washington). These communities will be our chief focus because there would be more room to work with numerically.

For the time series from Eugene, Longview and Whatcom, they passed the Augmented Dickey-Fuller (ADF) test. Therefore, they are stationary time series, we can and will use auto-vector regression to model them.

III. AUTO-VECTOR REGRESSION ANALYSIS

Eugene Daily testing

We use the first two years' data as training data and the last half year data as test data.

By using model selection process in regression analysis on the training data, we found that the following model is a good fit:

$$y_t = 20.9982 + 0.11361 y_{t-1} + 0.08089 y_{t-3} - 0.07221 y_{t-4} + 0.09033 y_{t-6} + 0.26747 y_{t-7}$$

Applying this prediction equation to the test data, we found the error measure is 17.2%, where the error measure is the average of every day's relative error in a given time period.

Longview Daily testing

By using the model selection process in regression analysis, we built the following model:

$$y_t = 1.46825 + 0.27378 y_{t-1} + 0.07552 y_{t-4} + 0.14003 y_{t-7} + 0.06939 y_{t-9} + 0.1166 y_{t-10} + 0.13172 y_{t-14} + 0.06247 * y_{t-15} - 0.10132 * y_{t-16} + 0.15136 * y_{t-20} + 0.07364 * y_{t-21} + 0.07221 * y_{t-23} - 0.09479 * y_{t-24}$$

Since here the window size is 24, to better test our model, we use the last 240 days' data as test data, and the first 642 days' data as training data.

When testing this prediction by the test data, we found that the error measure defined as above is 29.5%, which is a little bit high. This suggests that auto-regression model may not be a good enough model here.

Whatcom Daily testing

Again, we use the first two year data as training data, and the last half year data as test data.

By using the model selection process in regression analysis, we built the following model:

$$y_t = 39.89172 + 0.25875 y_{t-1} - 0.07478 y_{t-2} + 0.0628 y_{t-3} + 0.15060 y_{t-7}$$

By testing this prediction equation in the test data, we found that the error measure defined as above is 13.2%, which is low.

We continue the auto-vector regression modeling exploration. We separate the daily test volumes to female labs and male labs. Then we build the models for female labs and male labs for each region: Eugene, Longview and Whatcom. While the results were improved for Longview, the results were not improved for Eugene and Whatcom. In the following modeling, we use the first 700 days' data as training data and the last 182 days' data as test data.

In the following modeling, we use the first 700 days' data as training data and the last 182 days' data as test data.

Eugene Female Daily testing

The model is

$$y_t = 9.79 + 0.09884 y_{t-1} + 0.05657 y_{t-3} + 0.05462 y_{t-6} + 0.09846 y_{t-7} + 0.11428 y_{t-8} - 0.12392 y_{t-9} + 0.09418 y_{t-13} + 0.12659 y_{t-14}$$

The average relative error is 20.83%.

Eugene Male Daily testing

The model is

$$y_t = 11.60302 + 0.0874 y_{t-1} + 0.10141 y_{t-3} + 0.09598 y_{t-7} - 0.05759 y_{t-9} + 0.07597 y_{t-14} + 0.06801 y_{t-15} - 0.06720 y_{t-16} - 0.07214 y_{t-19} + 0.17427 y_{t-28}$$

The average relative error is 23.73%.

Longview Female Daily testing

The model is as follows

$$y_t = 1.9365 + 0.3245 y_{t-1} + 0.05393 y_{t-3} + 0.19137 y_{t-4} + 0.11276 y_{t-6} + 0.19913 y_{t-7}$$

The average relative error is 26.2%.

Longview Male Daily testing

The model is as follows

$$y_t = 0.141620 + 0.25970 y_{t-1} + 0.07819 y_{t-4} + 0.09792 y_{t-6} + 0.08313 y_{t-7} + 0.07499\, y_{t-9} + 0.09512 y_{t-11} + 0.06913 y_{t-13} + 0.14581 y_{t-14}$$

The average relative error is 29.08%.

Whatcom Female Daily testing

The model is as follows

$$y_t = 11.41537 + 0.22317 y_{t-1} + 0.08559 y_{t-7} + 0.09101 y_{t-13} + 0.06602 y_{t-18} + 0.09672 y_{t-21} + 0.10204 y_{t-28}$$

The average relative error is 23.25%.

Whatcom Male Daily testing

The model is

$$y_t = 28.9241 + 0.20425 y_{t-1} - 0.05252 y_{t-5} - 0.038647 y_{t-7}$$

The average relative error is 22.22%.

Daily Models

In the following we stratify the data to several groups. Namely, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday. For each we have one auto-vector regression model. So we will have 7 models for each region. We can see from the following results, the prediction results are improved. As before, we use the last two years data as training data and the last half year data as test data.

	Eugene	Whatcom	Longview
M	0.145	0.162	0.197
T	0.189	0.164	0.173
W	0.171	0.169	0.204
Th	0.128	0.192	0.274
F	0.145	0.158	0.295
S	0.142	0.222	0.192
Su	0.123	0.189	0.221
Average	0.149	0.189	0.222

IV. ADAPTIVE MODELING

A. *Cycles*

From previous analysis and modeling, we can easily see that the data from 1, 7, 14, 28 days play an important role. We would like to explore which 'day cycle' length was best for predicting demand for troponin lab tests. To clarify let's use an example.

Let the data be from 2 weeks as follows:

M: 1 T: 2 W: 4 TH: 7 F: 4 S: 2 SU: 6
M: 6 T: 3 W: 7 TH: 1 F: 4 S: 5 SU: 9

A 1 day cycle would mean we would compare Monday to Tuesday and then Tuesday to Wednesday and so on, a 7 day cycle would mean we compare Monday to next week's Monday. Let N be size of the input and n to be cycle size. Note that if n does not divide N we drop off the leftover values. We computed the average absolute error defined as $\sum_{i=1}^{n}\sum_{j=1}^{N}\left|\frac{X_{n+i+j}-X_{i+j}}{N}\right|$
for every n-cycle for n = 1 ... 30 and found the 5 minimum cycles to proceed with as our best.

We allowed for shifting of the start of the cycles in the computation. The reasoning is clear for the example of a 30-cycle, if the data started on the 15th of the month computing the 30-cycle starting from that day would not make as much sense as starting from the 1st of the month and comparing it with next month.

An example computation for a 7-cycle with no shift would be $\frac{|1-6|+|2-3|+|4-7|+|7-1|+|4-4|+|2-5|+|6-9|}{7} = 3$. A 6-cycle with a shift of 2 days would be $\frac{|4-3|+|7-7|+|4-1|+|2-4|+|6-5|+|6-9|}{6} = \frac{10}{6}$. The result is not surprising, the best cycles naturally are multiples of 7 and between days.

The result format is as follows: day cycle, start shift, error:

	Day Cycle	Start Shift	Error
Eugene	28	14	0.075
	7	1	0.076
	14	0	0.077
	21	20	0.078
	20	14	0.083
Whatcom	1	0	0.122
	14	13	0.128
	7	4	0.13
	28	0	0.122
	21	8	0.130
Longview	1	0	0.091
	7	1	0.096
	21	3	0.096
	14	1	0.096
	20	3	0.097

So to summarize the five best cycles are 1 day, 7 days, 14 days, 21 days, 28 days although there is a bit of variance between communities. For example in Eugene 1 day cycles does not even make it in the top 5 cycles.

In the following modeling, we will build our models based on these cycles.

B. Adaptive Modeling

We begin by building an adaptive auto-vector regression model for Eugene, Whatcom, and Longview, respectively. We build these models incrementally for each 28 days. That is to say we train our model to look at 28 days and then predict the 29th day. So for month 1 we would train our models with the samples

$x_1\ x_2\ \dots\quad x_{28}\ |\ Y_{29} = x_{29}$
$x_2\ x_3\ \dots\quad x_{29}\ |\ Y_{30} = x_{30}$
....
$x_{28}\ x_{29}\ \dots\ x_{56}\ |\ Y_{57} = x_{57}$

Then we would use this model to predict the 3rd month (28 days). More specifically to predict this month the model takes as input 28 days and outputs its estimate $\hat{Y_i}$

$x_{29}\ x_{30}\ \dots\quad x_{57}\ |\ \hat{Y}_{58}$
$x_{30}\ x_{31}\ \dots\quad x_{58}\ |\ \hat{Y}_{59}$
....
$x_{56}\ x_{57}\ \dots\quad x_{85}\ |\ \hat{Y}_{86}$

We calculate the Relative Error between the Y_i values and $\hat{Y}_i$ values where Relative Error is defined as $\sum_{i=1}^{n}\frac{|Y_i-\hat{Y}_i|}{Y_i}$. Since we are computing the error each time we incrementally build the model, the error is lessening as we add more and more data. Some results are:

Average relative error	Eugene	Whatcom	Longview
Predicting month 3	0.405	0.280	2.136
Predicting month 4	0.314	0.196	0.436
Predicting final month	0.125	0.124	0.188
Predicting last 6 months	0.145	0.157	0.198

C. Adaptive Modeling – Daily Modeling

Continuing with our adaptive model we want to see how well our model performs on each day. That is to say we train our model in precisely the same way as we previously had, however we only calculate the error in respect to a single day (eg. Monday).

	Eugene	Whatcom	Longview
M	0.123	0.142	0.183
T	0.09	0.096	0.221
W	0.106	0.111	0.168
Th	0.053	0.106	0.190
F	0.087	0.184	0.253
S	0.182	0.104	0.273

Su	0.29	0.124	0.028
Average	0.124	0.124	0.188

D. Adaptive Modeling –Sliding Rules

Currently our model uses all historic data in order to predict new days. Two main questions must be asked about this approach: "Can we remove very old historical data without affecting our model significantly?" and even more pessimistically "After a significant amount of time can historical data actually skew our future estimates?". If the first case is so we may want to remove old historical data for the sake of efficiency and a simpler model. If the second case is true we definitely want to remove old historical data from our model because it would actually improve our accuracy.

We explore these theories by testing two new models, a sliding rule maintaining 60 weeks' worth of historical data and a sliding rule maintaining 30 weeks of historical data. What we mean by sliding rule is that our model will only be fitted to the prescribed number of weeks of data so older data will be removed as we train on new data.

	Eugene	Whatcom	Longview
30-week	0.129	0.127	0.173
60-week	0.130	0.122	0.185
Full data	0.125	0.124	0.188
30-week average 6 months	0.152	0.162	0.206
60-week average 6 months	0.148	0.158	0.200
Full data average 6 months	0.145	0.157	0.198

E. Daily Modeling – Sliding Rules

In this subsection, we propose to use a sliding rule model of 104 and 52 days to predict future months. That is as time progresses older data is removed from the model as new data comes in so there is always 104 or 52 days of training data in our models (104 Mondays/Tuesdays/etc for each model). As in previous stratified models this is the prediction error over the last 6 months.

104 days	Eugene	Whatcom	Longview
M	0.139	0.161	0.188
T	0.186	0.176	0.183
W	0.172	0.170	0.190
Th	0.129	0.153	0.260
F	0.130	0.139	0.327
S	0.155	0.208	0.217
Su	0.175	0.225	0.208
Average	0.155	0.176	0.225

52 days	Eugene	Whatcom	Longview
M	0.185	0.244	0.220
T	0.237	0.275	0.256
W	0.210	0.210	0.239
Th	0.190	0.200	0.370
F	0.174	0.171	0.383
S	0.169	0.223	0.281
Su	0.272	0.247	0.306
Average	0.205	0.224	0.294

D. Confidence Interval

We have shown that our models are accurate in forecasting the demand for troponin lab tests. Now let us look at the confidence interval of the predictions.

To this end we define σ_n to be the standard deviation of our estimator up to the month we are predicting. Since we assume our predictions are unbiased we calculate σ_n by taking the square root of the MSE of our estimator where $MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2$. Our confidence interval is defined as: $x_{i1} = \hat{y}_i - 1.96\sigma$, $x_{i2} = \hat{y}_i + 1.96\sigma$.

The results on the 420 day sliding rule model with 28 day update schedule are:

	Eugene	Whatcom	Longview
Percent of values within confident interval	99.23%	98.60%	97.83%
Average length of confidence interval	44.50	67.50	43.75

V. Discussions and Conclusions

In this paper, the authors propose a computational approach to predicting the demand for future clinical services. Historical data, when analyzed using various modeling approaches, can proactively inform health systems regarding expected demand for their high cost services. These approaches may form the basis for evidence-based staffing strategies as well as possible community-centric approaches to proactively managing demand for health services. Finally, access to models with insight into the ebb and flow of group demand for health services will help inform new theories underlying the principles of individual demand for health services as mediated through various high yield human behaviors.

Acknowledgment

References

[1] G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.