Mathematical Approaches to Predictive Health Monitoring for Heart Failure Patients

Yi-Ju Chao
127 Vincent Hall, Minnesota Center of Industrial Mathematics

Abstract

This article introduces both statistical and the theoretical approaches to explore parameters, which can be used to predict the health of heart failure patients and to classify heart failure patients. From the statistical point of view, the comparison of covariance analysis and factor analysis with the patients' medical history suggests several parameters. From the theoretical point of view, a stochastic model with seven states was developed to describe the blood circulation system. Some covariance analysis of the theoretical model shows consistency with the empirical result.
Keywords: Heart failure; seven-state model; cardiovascular system; covariance analysis; factor analysis.

1 Introduction

In 1997 Jay N. Cohn M.D., president of Heart Failure Society of America[1] said "In the United States, heart failure is the number one reason people aged 65 years and older are hospitalized....[heart failure is] a growing problem in our society." The clinical syndrome of heart failure consists of shortness of breath, pitting edema, enlarged tender liver, engorged neck veins, and pulmonary rails. Generally speaking, heart failure is a condition in which the heart cannot pump enough blood to meet the metabolic needs of body tissues. Many patients develop heart failure from more than one cause. Some of the symptoms linked to heart failure are caused by the malfunction of organs other than the heart, especially the lungs, kidneys, and liver. Heart failure is closely linked to many forms of heart disease and is commonly diagnosed only after the diagnosis of heart disease. The treatment for heart failure commonly involves varying the workload of the heart by giving certain drugs, salt-free diets, or surgery.

This article introduces mathematical approaches to explore parameters which can be used to predict the health of heart failure patients and to classify heart

[1] http://www.hfsa.org

failure patients. There are two aspects of this project: one is the statistical analysis of the variables taken from heart failure patients, the other is the theoretical analysis of the cardiovascular system. From the statistical point of view, the observed variables are filtered according to the results produced by the Fast Fourier Transform (FFT) [12]. The comparison of covariance analysis and factor analysis on the filtered data with the patients' health history suggests several parameters which can be used to monitor heart-failure patients' health. From the theoretical aspect, we began by studying the literature to develop a stochastic model with seven states. The analysis of certain parameters from the seven-state stochastic model shows consistency with the statistical approach.

The observed data was acquired from the "Chronic Data Project"[2] gathered by Medtronic Inc.[3] The data were taken from twenty-one heart failure patients which are identified as U1-U6, P1-P9, and K1-K6 from 1995 to 1997. The patients were receiving vasodilators, diuretics, and inotropic agents. Twelve variables were from patients' right ventricles via voltage sensors, pressure transducers and oxygen sensors.

The twelve observed variables and the medical definitions are listed below.
(1) Heart rate (bmp): The number of heart beats per minute.
(2) Systolic pressure (mmHg): The medical definition is the maximal blood pressure occurring during the contraction of the left ventricle, in the case of this project, experimental data is the maximal blood pressure of the right ventricle.
(3) Diastolic pressure (mmHg): The medical definition is the minimal blood pressure occurring during the contraction of the left ventricle, in the case of this project, experimental data was taken at the end of the diastole of the right ventricle.
(4) Pulse pressure (mmHg): Pulse pressure is the difference between systolic pressure and diastolic pressure.
(5) Max. positive dP/dt: The maximal slope of the pressure curve over each cycle.
(6) Max. negative dP/dt: The minimal slope of the pressure curve over each cycle.
(7) ePAD (mmHg): An estimate of pulmonary artery pressure. The pressure was taken at the time when Max. positive dP/dt occurs.
(8) Pre-ejection interval (ms): The medical definition is the time between the beginning of systole and the opening of aortic valve. The experimental data records the time between the beginning of systole and the time when Max. positive dP/dt occurs.
(9) Systolic time interval (Sti.) (ms): The time between systole.
(10) Oxygen saturation: The experimental oxygen saturation occurring in the right ventricle.
(11) Activity (count): A measure of patients' activity.

[2]The Division of Heart Failure Management
[3]7000 Central Avenue N.E., Minneapolis, MN55432

(12) Ptemp: An estimate of body temperature.

This project lasted for three months, from June/1997 to Sep./1997. Due to the short time frame, Attention was selected patients (P1, P5, P7, U2, U3, U5) and several variables (heart rate, pressure data, oxygen saturation). A systematic comparison of the results produced by the statistical analysis of P7 with U5 suggested promising directions for further studies. The computer software SAS and Matlab 5.0 tool box (including Signal processes) were used to compute required parameters.

This article is organized as follows. Section 1 introduces this project. A description of how the observed data was filtered is in section 2. Analysis techniques, covariance analysis and factor analysis of the filtered data, and the classification of patients are discussed in section 3. A development of the seven-state model is discussed section 4. Section 5 discuss the model analysis, in which the covariance between certain variables is derived according to the appropriate choices of control mechanisms of the cardiovascular system. Parameters, classification, and summary are discussed in section 6. Section 7 contains suggestions for further research.

2 Filtering

Filtering the data can make the data easier to examine. By examining plots of the filtered data, the appropriate analysis techniques may be chosen. The first was to take the FFT [16] and to select the appropriate filter accordingly.

The FFT is a tool used to analyze the spectrum of signal processes. The sample, taken every four hours, is the median of the raw data. Fig.1. is an example of the Fourier analysis results. The upper graph is the curve of diastolic pressure against time with a sample period of four hours for the patient U3. The lower curve represents the power spectrum of the upper curve. The lower curve is symmetric with respect to the center line. The obvious high peaks represent one-day and half-day cycles respectively. Also there are some observed strong low frequency components appearing on the graph. These are observed for each variable (heart rate, systolic pressure, diastolic pressure, pulse pressure, Max. positive dP/dt, Max. negative dP/dt, ePAD, oxygen saturation, activity, pre-ejection time interval, systolic time interval, ptemp) from patients (P1, P5, P7, U2, U3, U5). These results suggest the use of a low-pass filter to capture the low frequency components of the signals and attenuate others.

Figure.1. Raw Data and Fourier Analysis for Patient U3

The low-pass filter is as follows. Let $u(t)$ represent the observed data and $y(t)$ represent the filtered data. The system is assumed to be linear, time-invariant, and satisfying

$$\mu y'(t) + y(t) = u(t) \text{ for constant } \mu,$$

where $y'(t)$ denotes the first derivative with respect to t. The relation between $y(t)$ and $u(t)$ is

$$y(t) = \frac{1}{\mu} \int_0^t u(s) e^{-\frac{t-s}{\mu}} ds.$$

In a discrete case, let $u(n)$, with n integer, be the n-th observation, and let $y(n)$ to be the filtered data. With the same assumptions, the relation between $u(n)$ and $y(n)$ is as follows

$$y(n) = \frac{1}{\mu} \sum_{i=1}^{n} u(i) e^{-\frac{n-i}{\mu}}.$$

The sample period was one-half hour. Missing data were filled by linear interpolation.

The constant μ affects the range of the frequency, phase-shift, and amplitude-change of the filtering[12]. This filter used μ to be 50, which made the filtered curves easy to observe.

3 Analysis of the Filtered Data

Covariance analysis and Factor analysis were used to investigate the pattern of variables and to classify patients. Covariance analysis is a technique to investigate the deviation of a variable and the correlation among variables quantatively. Factor analysis is another tool to analyze systems. By using this technique, we can find hypothetical variables, or factors, which contain the essential information of systems. The covariance analysis and factor analysis on filtered data compared to the patients' medical history suggested some potential parameters of monitoring the health of heart failure patients or classifying patients. Patients P7 and U5 were selected as examples.

3.1 Covariance Analysis

Assume two signal processes $X = (x_1, x_2, \cdots, x_N)$ and $Y = (y_1, y_2, \cdots, y_N)$, the covariance of the signals X and Y is defined by

$$cov(X,Y) = \sum_{i=1}^{N} \frac{(x_i - \bar{x})(y_i - \bar{y})}{N} = \sum_{i=1}^{N} \frac{x_i y_i - \bar{x}\bar{y}}{N}, \tag{3.1}$$

where $\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N}$ is the mean of the signal X, and $\bar{y} = \frac{\sum_{i=1}^{N} y_i}{N}$ is the mean of the signal Y. The normalization of $cov(X,Y)$ is defined by

$$corr(X,Y) = \frac{cov(X,Y)}{\sqrt{cov(X,X)cov(Y,Y)}},$$

which is called the correlation coefficient of X and Y. Notice that $corr(X,Y)$ is between -1 and 1. If there are M signal processes, say $S_1, S_2, \cdots, S_M$, we shall get a $M \times M$-matrix $(corr(S_i, S_j))_{i=1,\ldots,M;j=1,\ldots,M}$ which is called the correlation matrix. If the correlation coefficient is positive, the two signals are more likely to be rising and dropping simultaneously. On the other hand, if the correlation coefficient is negative, the two signals are more likely to be rising and dropping oppositely. Fig.2. shows the motivation to use "windowed" covariance analysis. For patient P7, the curves seem to be stable during interim, and dropping/rising dramatically during hospitalization. In addition, the oxygen saturation and the systolic pressure are rising during hospitalization, while the diastolic pressure were dropping. For patient U5 (see Fig.3.), it is not possible to tell the trends during hospitalization from the graphs. The windowed covariance analysis observed the correlation matrices of different periods of time.

H1 stands for hospitalization.H2 stands for hospitalization.D stands for death of patients.

Figure.2. Covariance Analysis for Patient P7

H stands for hospitalization.D stands for death of patients.

Figure.3. Covariance Analysis for Patient U5

According to patient P7's medical record, the data is separated into ten periods of time, which are interim(1) (I.1), eight days before hospitalization(1) (B.H.1), hospitalization(1) (H.1), eight days after hospitalization(1) (A.H.1), interim(2) (I.2), eight days before hospitalization(2) (B.H.2), hospitalization(2) (H.2), eight days after hospitalization(2) (A.H.2), interim(3) (I.3), and death of patient (D). The reason of choosing eight days before/after hospitalization is that the plots start to rise/drop approximately eight days before/after hospitalization. The correlation between oxygen saturation and the one between

Table.1. Oxygen Saturation Row of Correlation Matrices for Patient P7

O.S.[1]v.s.	H.R.[2]	Sys.P.[3]	Dias.P.[4]	Pulse P.[5]	Pos.dP/dt[6]	Neg.dP/dt[7]	ePAD
I.1	0.0566	0.6309	-0.5411	0.7633	0.7030	0.6587	0.0438
B.H.1	-0.4798	0.6859	-0.8503	0.6521	0.8681	0.8313	0.0497
H.1	-0.8945	0.7881	-0.9801	0.8252	0.6518	0.5658	0.3763
A.H.1	0.6539	-0.4043	-0.1083	-0.5736	0.3373	-0.0357	0.1935
I.2	0.5514	-0.4512	-0.8686	-0.2249	0.5013	-0.3881	0.7304
B.H.2	-0.6454	0.8078	-0.3868	0.8798	0.7929	0.8676	0.6530
H.2	-0.8775	0.8391	-0.9165	0.9286	0.8926	0.9205	0.1169
A.H.2	0.0415	-0.0753	-0.8720	0.0357	0.1069	-0.4182	0.7561
I.3	-0.5046	0.0157	-0.8055	0.3707	0.1658	0.3576	0.2918
D	-0.0210	0.7897	-0.8524	0.7763	0.7460	0.7433	0.3603

Table.2. Oxygen Saturation Row of Correlation Matrices for patient U5

O.S.[1]v.s.	H.R.[2]	Sys.P.[3]	Dias.P.[4]	Pulse P.[5]	Pos.dP/dt[6]	Neg.dP/dt[7]	ePAD
I	-0.7164	0.4905	-0.6681	0.5872	0.5434	0.5575	0.1747
B.H.	0.1990	-0.2610	-0.4900	0.5774	0.5276	0.7246	-0.6483
H.	-0.2663	0.7369	-0.5666	0.9192	0.8983	0.9006	0.0324
A.H.	-0.0910	0.0751	-0.3638	0.2032	-0.1761	0.2416	-0.2743

pressure data are given on Table.1. and Table.2. respectively.

Notice that in Table.1., the correlation coefficient $corr(O.S., H.R.)$ is negative right before and at hospitalization, and positive after, while the correlation coefficient $corr(O.S., Dias.P.)$ is always negative.

For patient U5 (see Table.2.), the data is seperated into interim (I), seven days before hospitalization (B.H.), hospitalization (H), and after hospitalization (A.H.). The correlation coefficient $corr(O.S., H.R.)$ is positive as the patient became sick, negative otherwise. This result is opposite to the result of patient P7. The correlation coefficient $corr(O.S., Dias.P.)$ is still always negative. A seven-state model is developed to explain this covariance result. The sign of $corr(O.S., Dias.P.)$ can be derived from the model (see section 4).

[1] O.S. stands for oxygen saturation.
[2] H.R. stands for heart rate.
[3] Sys. P. stands for the systolic pressure.
[4] Dias. P. stands for the diastolic pressure.
[5] Pulse P. stands for the pulse pressure.
[6] Pov.dP/dt stands for the max. positive dP/dt.
[7] Neg.dP/dt stands for the max. negative dP/dt.

3.2 Factor Analysis

The goal of factor analysis is to construct a few hypothetical variables, or factors, which are supposed to contain the information of the observed variables and how strong each factor is. The basic assumption is that each factor is a linear combination of the observed variables. In practice, the factors and the magnitude of the factors, which shows how strong factors are, can be gotten by the singular value decomposition (SVD) [13].

Assume that there are m variables and there are n observations for each variable. The data matrix can be denoted by a $m \times n$-matrix $A = (a_{ij})_{1 \le i \le m; 1 \le j \le n}$. According to the Eckard-Young Theorem [3], there exists a positive integer $p \le m$, a $n \times p$-matrix U with orthonormal columns, a $m \times p$-matrix V with orthonormal columns, and a matrix

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 & 0 \\ 0 & \sigma_2 & & & 0 \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & & & \sigma_{p-1} & 0 \\ 0 & 0 & \cdots & 0 & \sigma_p \end{pmatrix}$$

with $\sigma_1 \ge \cdots \ge \sigma_p$ such that

$$A = U \Sigma V^{'}.$$

This is called the singular value decomposition of the matrix A.

The matrix U is the coefficient matrix. Let $V = (v_1, \cdots, v_p)$, where v_i are the column vectors of V for $1 \le i \le p$. The matrix $\Sigma V^{'} = (\sigma_1 v_1, \cdots, \sigma_p v_p)^{'}$ is the factor matrix, where σ_i are called singular values and v_i are called singular vectors. Since V is an orthonormal matrix, these factors $\sigma_i v_i$ are orthogonal with each other and have magnitude σ_i. The singular value σ_i shows how strong the factor $\sigma_i v_i$ is. The components of the singular vector v_i, a m-column vector, shows the contribution of each variable to this factor. In practice, the i-th factor is significant[4] if

$$\frac{\sigma_1}{\sigma_i} \le \sqrt{3m}, \text{ for } 2 \le i \le m. \tag{3.2}$$

Notice that the factors with bigger singular values are stronger factors. The big singular values and the corresponding singular vectors well determine the information of the system.

For consistency, the data was separated into different periods of time in covariance analysis. Table.3., Table.4., and Fig.4. are the singular values, the first singular vector, and the ratio of the first two singular values respectively for patient P7. Table.5., Table.6., and Fig.5. show the corresponding results for patient U5. Since the units of each of the variables are different, the data are normalized by dividing by the maximum of each variable.

[4] Numerical experiments with random matrices show consistency with this criterion.

Table.3. Singular Values for Patient P7

	I.1	B.H.1	H.1	A.H.1	I.2	B.H.2	H.2	A.H.2	I.3	D
σ_1	121.36	59.97	61.75	62.20	105.44	59.40	40.67	62.76	120.75	82.95
σ_2	7.57	5.96	9.56	2.52	6.30	5.19	4.11	4.83	6.69	6.49
σ_3	6.80	3.12	2.67	1.32	5.60	2.84	4.09	1.75	1.36	0.75
⋮[8]	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table.4. First Singular Vector for Patient P7

	I.1	B.H.1	H.1	A.H.1	I.2	B.H.2	H.2	A.H.2	I.3	D
H.R.	0.250	0.267	0.237	0.295	0.273	0.269	0.256	0.291	0.253	0.250
Sys. P.	0.313	0.312	0.314	0.291	0.303	0.295	0.316	0.311	0.307	0.313
Dias. P.	0.239	0.248	0.231	0.285	0.194	0.263	0.284	0.254	0.278	0.239
Pulse P.	0.291	0.277	0.279	0.292	0.301	0.280	0.289	0.311	0.309	0.291
Pos.dP/dt	0.283	0.290	0.253	0.270	0.302	0.296	0.289	0.303	0.281	0.283
Neg.dP/dt	0.304	0.318	0.303	0.302	0.302	0.298	0.299	0.310	0.307	0.304
ePAD	0.311	0.310	0.329	0.303	0.285	0.308	0.314	0.296	0.298	0.311
O.S.	0.277	0.292	0.300	0.285	0.300	0.282	0.278	0.256	0.286	0.277
Act[9]	0.200	0.215	0.139	0.220	0.228	0.236	0.133	0.154	0.143	0.201
Pei[10]	0.326	0.306	0.336	0.304	0.316	0.311	0.298	0.315	0.324	0.326
Sti[11]	0.311	0.290	0.331	0.302	0.309	0.291	0.329	0.311	0.321	0.311
ptemp	0.331	0.321	0.344	0.305	0.324	0.324	0.326	0.311	0.312	0.331

[8]Since σ_i are descending, the value of σ_i, for $3 < i$, are relatively small.
[9]Act stands for the activity of the patient.
[10]Pei stands for the pre-ejection time interval.
[11]Sti stands for the systolic time interval.

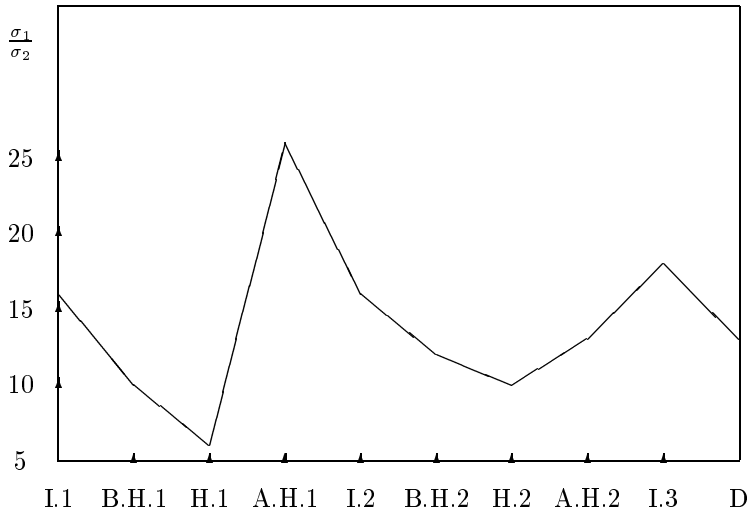

Figure.4. $\frac{\sigma_1}{\sigma_2}$ for Patient P7

The factor analysis conclusion for patient P7 is as follows. Table.3. suggests that the system has primarily one significant factor according to the criterion (3.2). Since there is one significant factor in this system, the first singular vector shows the significant contribution of each variable. Table.4. shows that the components of the first singular vector are pretty even, which means there is no prominent variable in the system. Fig.4. shows the curve of the parameter $\frac{\sigma_1}{\sigma_2}$, which is the ratio of the first singular value and the second singular, against different time intervals. Decaying $\frac{\sigma_1}{\sigma_2}$ is correlated with declining health. The last fact can be interpreted to mean that the second factor shows stronger effect when the patient's health is declining.

The conclusion of factor analysis for patient U5 is as follows. According to the criterion (3.2), Table.5. also shows that the system has primarily one significant factor. Table.6. shows that four dominant variables, which are maximal positive dP/dt (Pos. dP/dt), maximal negative dP/dt (Neg. dP/dt), Sti. , and Ptemp, contribute to the significant factor. Fig.5. suggests that the increasing $\frac{\sigma_1}{\sigma_2}$ is correlated with declining health. This shows that the second factor is weaker as the patient's health is declining, which is opposite to the result for patient P7.

Table.5. Singular Values for Patient U5

	I	B.H.	H.	A.H.
σ_1	32572.4	19567.26	8212.26	8443.62
σ_2	1519.4	764.38	88.59	63.50
σ_3	379.6	276.05	58.09	28.16
σ_4	202.1	201.92	21.80	10.69
$\vdots^8$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table.6. First Singular Vector for Patient U5

	I	B.H.	H.	A.H.
H.R.	0.001087	0.001415	0.00157	0.00146
Sys. P.	0.000097	0.000142	0.00157	0.00152
Dias. P.	0.000089	0.000119	0.00158	0.00143
Pulse P.	0.000089	0.134450	0.00152	0.00144
Pos. dP/dt	0.6661	0.5931	0.5640	0.5787
Neg. dP/dt	0.6018	0.6079	0.5981	0.5961
ePAD	0.02011	0.025695	0.0354	0.03499
O.S.	0.066303	0.074913	0.0755	0.07502
Act[9]	0.001315	0.000030	0.00017	0.00007
Pei[10]	0.008111	0.1192	0.14820	0.15966
Sti[11]	0.3752	0.4331	0.46265	0.4611
ptemp	0.2049	0.2683	0.2849	0.2605

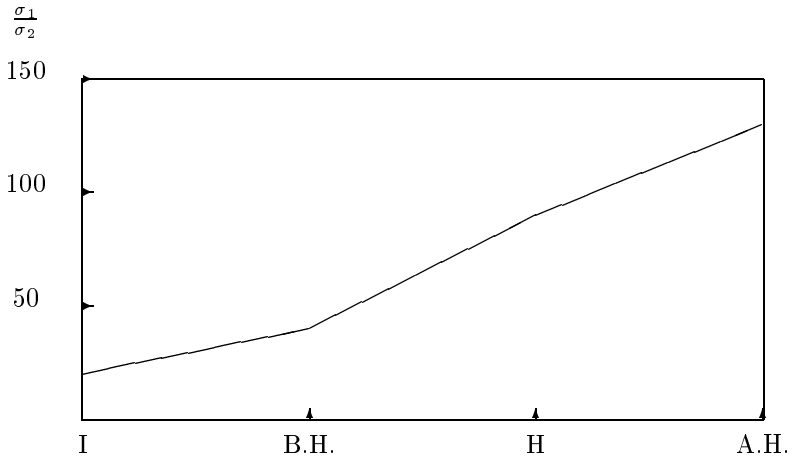

Figure.5. $\frac{\sigma_1}{\sigma_2}$ for Patient U5

4 Seven-state Model

In this section, a model describing the heart and blood circulation in the body is discussed to explain the result produced by covariance analysis. The function of the heart is to pump blood. The blood carries oxygen from the lungs to the various tissues of the body and carries carbon dioxide from these tissues back to the lungs. Since the circulation of blood forms a closed loop, the description can begin anywhere. The right ventricle pumps blood into the pulmonary arteries which distribute blood to the tissues of the lungs. The blood going through the lungs is oxygenated on the way to the left side of the heart. The left ventricle pumps blood rich in oxygen into the systemic arteries to meet the oxygen consumption need of the body. The blood from the body goes back to the right heart. This completes the circulation. The blood circulation is modeled as a four-site mechanical system. Fig.6. is a schematic description of this mechanical system. The arrows show the direction of the blood circulation. The right side of the heart (R.H.), the lungs (L.), the left side of the heart (L.H.), and the body (B.) are represented by the little squares. The points of the blood circulation entering into the R.H., the L., the L.H., and the B. are designated by site 1, site 2, site 3, and site 4, respectively.

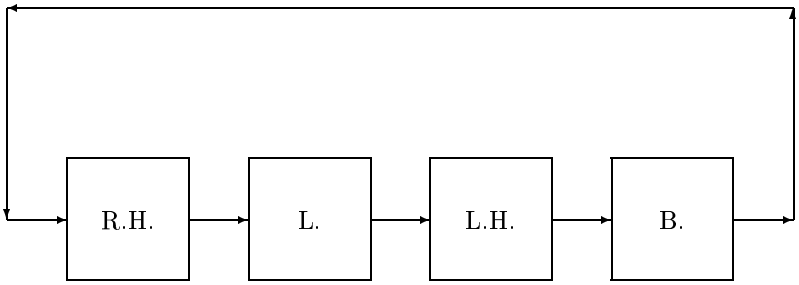

Figure.6. Blood Circulation System

Some important physical variables are needed in a quantative description of the circulation. Since one of the functions of the circulation is to deliver oxygen through the body, the oxygen saturation at each site should be taken into consideration. The oxygen saturations at each site are denoted by O_1, O_2, O_3, O_4 respectively. Since there is less oxygen consumption and less oxygen saturating the blood on the route from the right side of the heart (or site 1) to the lungs (or site 2), the oxygen saturation at site 1, O_1, is assumed to be equal to the oxygen saturation at site 2, O_2. Similarly, the oxygen saturation at site 3 and the one at site 4 are also close, so O_3 is assumed to be equal to O_4. Experimental data shows that O_3 is usually pretty consistent and approximately 95% [7], so O_3 is assumed to be a parameter in the model. The observed oxygen saturation should be O_1, because it is taken at the right ventricle. The heart rate, the number of heart beats per unit time, is denoted by H. The Q is the volume of blood per unit time passing a point in the circulation. The pressures at each site are denoted by P_1, P_2, P_3, P_4 respectively.

Some assumptions are necessary in a steady-state model. The Q is assumed to be consistent at every point of the circulation. The pressures P_1, P_2, P_3, P_4 are all assumed to be consistent in time. In physiology, the pressures P_1, P_2 correspond to the diastolic pressure and systolic pressure of the right ventricle respectively, and P_3, P_4 correspond to the diastolic pressure and systolic pressure of the left ventricle respectively. The above physical variables are the internal states of the circulation. In fact, the circulation is affected by activity. The oxygen consumption rate of the body is controlled by activity, which is indicated by the oxygen consumption rate, denoted by O_c.

The model is composed of variables P_1, P_2, P_3, P_4, Q, H, O_1 and one input O_c. An assumption is the volume of the blood vessel and the blood flow rate is not changing with time, and the blood is incompressible, the pressure drop is assumed to be proportional to the blood flow rate. The mathematical relation is as follows

$$R_b = \frac{Q}{P_4 - P_1}, \tag{4.3}$$

where R_b is called the total peripheral resistance (TPR) which is the resistance

offered to the total output of the left side of the heart by the peripheral vascular bed of the body. Because the symmetry of form between the right and left sides of the heart and the systemic and pulmonary circulations, there is a similar equation for the pulmonary circulation,

$$R_l = \frac{Q}{P_2 - P_3}, \tag{4.4}$$

where R_l is the total resistance offered to the output of the right side of the heart by the peripheral vascular bed of the lungs. It may be argued that the linear assumption between blood-flow rate and pressure differences may be too idealistic, especially, since the volume of blood vessels changes with pressure differences. In fact, tissues exhibit reasonably constant values of resistances under conditions where the volume of blood vessels remain constant. The linear assumption is introduced for simplicity. The reader may find more detailed discussion of these formulas in [Ch.8 in [2]] and [Ch.5 in [1]]. Improved, more sophisticated formulas are discussed later on in this article.

Experimental and clinical observations suggest the so-called Frank-Starling mechanism. According to the Frank-Starling relation, the amount of energy generated by the ventricle in one beat is a function of filling pressure [Ch.1 in [4], Ch.4 in [1], and Ch.3 in [7]]. The energy is called stroke work. The filling pressure of the pulmonary circulation is approximately equal to the diastolic pressure of the right ventricle and the filling pressure of the systemic circulation is the diastolic pressure of the left ventricle. The stroke work of the right heart and that of the left heart can be represented by the functions $F_r(P_1)$ and $F_l(P_3)$ respectively. The power, which is the total work generated by the ventricle per unit time, is the product of the blood-flow rate and the pressure difference over each beat. The pressure difference is approximated by the difference of the systolic pressure and the diastolic pressure. The combination of the Frank-Starling mechanism and the above power equations implies the following two equalities for the right ventricle and the left ventricle respectively.

$$Q(P_2 - P_1) = F_r(P_1)H, \tag{4.5}$$

and

$$Q(p_4 - P_3) = F_l(P_3)H. \tag{4.6}$$

The curve of stroke work against filling pressure is called the Frank-Starling curve, which plays a very important role in clinical treatments. For example, a patient with heart failure due to predominant systolic pressure dysfunction tends to have higher filling pressure and lower stroke work, so the Frank-Starling curve tends to be lower than the one for normal subjects. Medicine such as diuretics, vasodilators, and inotropic agents, can move patients to a higher Frank-Starling curve [Ch.34 in [5]]. The Frank-Starling curves vary with individuals and their health conditions. The estimate of the Frank-Starling curves for heart-failure patients is an important subject.

One of the well-accepted methods to measure blood-flow rate is Fick's Principle [P.75 in [1]], [P.49 in [11]], [7]. This principle is an application of the law of conservation of mass. It is derived from the fact that the quantity of oxygen entering from the right side of the heart plus the quantity consumed by the body must equal the quantity of oxygen delivered by the lungs. The mathematical formula of Fick's principle is below.

$$O_c = KQ(O_3 - O_1), \tag{4.7}$$

where K is a constant.

The above five equations 4.3 - 4.7 constitute the so-called seven-state model. The system of equations are given as follows:

$$P_2 - P_3 = R_l Q$$

$$P_4 - P_1 = R_b Q$$

$$Q(P_2 - P_1) = F_r(P_1)H$$

$$Q(P_4 - P_3) = F_l(P_3)H$$

$$O_c = KQ(O_3 - O_1)$$

This is a model with seven states $P_1, P_2, P_3, P_4, O_1, H, Q$, and one input O_c. The parameters R_l, R_b, K and the functions F_r, and F_l depend on individual patients.

5 Model Analysis

5.1 Control Mechanisms of Cardiovascular System

The solutions of the seven-state model form a three-dimensional manifold because there are eight unknowns and five equations, therefore, three independent variables need to be chosen to parametrize the solution set. The independent variables should be chosen according to the control mechanisms which regulate the circulation and make it serve the needs of the body. The terms O_c, P_1 were chosen as independent variables for external control reasons, and P_4 was chosen as another independent variable for intrinsic control reasons.

Classically, three types of controls have been recognized to play major roles in the operation of the cardiovascular system [Ch.8 in [11]]. They are metabolic control, myogenic control, and neural control.

Metabolic control, which alters the chemical composition of intra- and extravascular fluid, affects the circulation [Ch.19 in [6]]. One important indicator of the metabolic rate is the oxygen consumption rate O_c, therefore it is reasonable to choose O_c as an independent variable. The oxygen consumption rate, which is controlled by activity, is an external control of the system.

Myogenic control affects the system by changing the resistances when the transmural pressure of the blood vessels changes. In the seven-state model, the resistances are assumed to be constant since the volume of blood vessels are assumed to remain consistent. A more accurate model would include state dependent resistances, but the approximation might be good enough for some purposes.

The other important factor is neural control. The baroreceptors, which are located at the carotid arteries and the arch of the aorta, transmit nerve impulses to the brain stem according to the detection of pressure P_4 [Ch.12 in [14]], [Ch.20 in [2]]. In this way, the cardiovascular system can be autoregulated, so the variable P_4 is chosen as an intrinsic control.

The circulation of a healthy person is predominately controlled by the above three mechanisms. Heart-failure patients are usually receiving medical treatment, such as diuretics, vasodilators, or inotropic agents. One goal of these medicines is to reduce the filling pressure, preload [Ch.13 in [9]]. Diuretics improve symptoms of heart failure by lowering the preload (or ventricular filling pressures) of the heart. Vasodilators increase the cardiac output while also reducing the preload. Inotropic agents can result in higher cardiac output for a given level of ventricular filling pressure. Combinations of these three kinds of drugs will often yield additive effects in hemodynamics [5]. So P_1 is taken as an independent variable.

The control mechanisms cooperate to regulate the cardiovascular system. The physical states from the heart are expected to achieve equilibrium by intrinsic and external control mechanisms. A large deviation from normal equilibrium is expected under abnormal conditions. To identify normal and abnormal conditions, one examines covariance analysis from the seven-state model by choosing independent variables O_c, P_1, and P_4.

5.2 Covariance Analysis from Model

5.2.1 Preliminary

This section introduces the necessary notation and theory [10] in the covariance analysis of the seven-state model.

Let $\xi(s)$ be a stochastic process and $f(t)$ be a real-valued function, the Itô integral is denoted by

$$\int_0^t f(s)d\xi(s),$$

where $d\xi_s$ is the stochastic differential. For two stochastic processes $\xi_1(s)$ and $\xi_2(s)$, the quadratic variational process corresponding to $\xi_1(s)$ and $\xi_2(s)$ is denoted by $< \xi_1, \xi_2 >_s$.

Numerically, the expectation of the quadratic variational process $E < \xi_1, \xi_2 >_t$

can be approximated by

$$\frac{\sum_{i=1}^{N}(\xi_1(i)\xi_2(i) - \bar{\xi_1}\bar{\xi_2})}{N},$$

where $\xi_j(i) = \xi_j(i\Delta t)$ with $\Delta t = \frac{t}{N}$ and

$$\bar{\xi_j} = \frac{\sum_{i=1}^{N}\xi_i}{N} \text{ for } j = 1, 2.$$

The following is the connection between (3.1) and the quadratic variational process

$$corr(\xi_1, \xi_2) \sim cov(\xi_1, \xi_2) \approx E < \xi_1, \xi_2 >_t,$$

where $A \sim B$ means that the numbers A and B have the same sign, and $A \approx B$ means that B can be approximated by A. In $corr(\xi_1, \xi_2)$ and $cov(\xi_1, \xi_2)$, the processes $\xi_j = (\xi_j(i))_{i=1}^{N}$ for $j = 1, 2$, are discrete. In $E < \xi_1, \xi_2 >_t$, the processes $\xi_j = \xi_j(t)$, for $j = 1, 2$, are continuous.

Lemma 5.1 *Let $u(x)$ be a real-valued smooth function defined on the d-dimensional Euclidean space E_d. Let ξ_s be a d-dimensional stochstic process that has a stochastic differential denoted by $d\xi_s$. Then the process $u(\xi_s)$ also has a stochastic differential and*

$$du(\xi_s) = \sum_{i=1}^{d} u_{x^i}(\xi_s)d\xi_s^i + \frac{1}{2}\sum_{i,j=1}^{d} u_{x^i x^j}(\xi_s)d\xi_s^i d\xi_s^j. \tag{5.8}$$

The above formula is called Itô's formula.

5.2.2 Derivation of Covariance

The mathematical interpretation of the regulation by control mechanisms is described next. Let X be the state space of the seven-state model with coordinates P_1, P_2, P_3, P_4, O_1, H, Q, O_c. Let M be the solution set which is a three-dimensional manifold. The independent variables O_c, P_4, and P_1 parametrize the manifold M. Assume that the body's regulation by control mechanisms keeps the physical variables near a certain equilibrium over a short time. The equilibrium, which is denoted by $\overline{x} = (\overline{P_1}, \overline{P_2}, \overline{P_3}, \overline{P_4}, \overline{O_1}, \overline{H}, \overline{Q}, \overline{O_c})$, should be a point of the manifold M.

In fact, the variables P_1, P_2, P_3, P_4, O_1, H, Q, O_c do change over time on account of unknown factors, which were taken to be random. A refined stochastic model was developed to approximate the effect. Assume a stochastic process $x(t) = (P_1(t), P_2(t), P_3(t), P_4(t), O_1(t), H(t), Q(t), O_c(t))$, with time t. Due to the control mechanisms, the process $x(t)$ is driven by the following independent noise processes

$$O_c(t) = \overline{O_c} + \int_0^t \alpha(s)dw_1(s),$$

$$P_4(t) = \overline{P_4} + \int_0^t \beta(s) dw_2(s),$$

and

$$P_1(t) = \overline{P_1} + \int_0^t \gamma(s) dw_3(s),$$

where w_1, w_2, w_3 are independent Wiener processes. Making a quasi-steady-state assumption, the stochastic process $x(t)$ tends to remain on the manifold M. Moreover, in the absence of any significant changing the control variables, the stochastic process $x(t)$ is expected to travel within a small neighborhood of a fixed equilibrium point.

The covariances $cov(O_1, Q)$ and $cov(O_1, P_1)$ were computed to test the predictive value of the model with the empirical result produced by covariance analysis of the data. Equation 4.3 and Fick's Principle (4.7) imply the equation

$$O_1 = O_3 - \frac{R_b O_c}{K(P_4 - P_1)}.$$

Applying Itô's formula to the above formula, the stochastic differential of O_1 is

$$dO_1 = -\frac{R_b}{K}\left(\frac{d(O_c)}{P_4 - P_1} - \frac{O_c dP_4}{(P_4 - P_1)^2} + \frac{O_c dP_1}{(P_4 - P_1)^2} + \frac{O_c \beta^2 ds}{(P_4 - P_1)^3} + \frac{O_c \gamma^2 ds}{(P_4 - P_1)^3}\right). \tag{5.9}$$

The direct calculation of equation 4.3 gives the stochastic differential of Q

$$dQ = \frac{dP_4 - dP_1}{R_b}. \tag{5.10}$$

Equation 5.9 and 5.10 imply the covariance of O_1 and Q to be

$$cov(O_1, Q) \approx E < O_1, Q >_t = E \int_0^t d < O_1, Q >_s,$$

with

$$d < O_1, Q >_s = \frac{2O_c(\alpha^2 + \beta^2)}{K(P_4 - P_1)^2} ds,$$

which is positive. The covariance of O_1 and P_1 is

$$cov(O_1, P_1) \approx E < O_1, P_1 >_t = E \int_0^t d < O_1, P_1 >_s,$$

with

$$d < O_1, P_1 >_s = -\frac{R_b O_c \gamma^2}{K(P_4 - P_1)^2} ds,$$

which is negative.

Table.7. The Similarity between P7 and U5

	Patient P7	Patient U5
Covariance Analysis	$Corr(O.S., Dias.P.) < 0$	$Corr(O.S., Dias.P.) < 0$
Factor Analysis	One-dim. system	One-dim. system

Table.8. The Difference between P7 and U5

	Patient P7	Petient U5
Covariance Analysis	$Corr(O.S., H.R.) < 0$ right before and while hospitalized, > 0 after.	$Corr(O.S., H.R.) > 0$ when sick and < 0 otherwise.
Factor Analysis	Decreasing $\frac{\sigma_1}{\sigma_2}$ correlated with declining health; No prominent variables.	Increasing $\frac{\sigma_1}{\sigma_2}$ correlate with declining health; Four dominant variables.

6 Summary

The covariance analysis and the factor analysis produced results that suggest several parameters to predict the health of heart-failure patients. The sign of the correlation coefficient $corr(O.S., H.R.)$ seems to be a consistent tool to monitor the health within each patient, but inconsistent across the patients. The factor analysis suggests the parameter $\frac{\sigma_1}{\sigma_2}$ may predict health, which is also consistent within each patient, but inconsistent across the patients.

The analysis also suggests some parameters to classify the patients. Table.7. and Table.8. display the similarity as well as the difference between patients P7 and U5 respectively.

A seven-state model was developed to explain the covariance analysis result. The correlation $Corr(O.S., Dias.P.)$ was derived from the model. The result produced by the covariance analysis is consistent with the theoretical result in that the sign of $Corr(O.S., Dias.P.)$ is always negative. The positive correlation between the oxygen saturation and the blood-flow rate is also derived from the model, which is consistent with the prediction. Efforts made to derive an accurate estimate of blood-flow rate from the data available were unsuccessful, so we cannot say whether the theoretical predictor is consistent with our data.

7 Suggestion for Further Research

7.1 Windowed Covariance Analysis

7.1.1 What is the appropriate time interval?

In the windowed covariance analysis, the time intervals were decided by observing the plots of the data. The way of determining the time intervals (e.g. How many days before/after the period of hospitalization?) can affect the covariance analysis result. Moreover, from the aspect of predicting patients' health, this should be an important subject. Sophisticated methods of checking the stability of the signal processes should be brought in, such as testing hypothesis, autocorrelation.

7.1.2 Is the correlation $corr(O.S., H.R.)$ criterion useful?

As mentioned in the summary, the correlation $corr(O.S., H.R.)$ seems to be a parameter for predicting the health of the heart failure patients. Deriving the correlation $corr(O.S., H.R.)$ from the seven-state model to support this criterion is another direction for further research. The correlation $corr(O.S., H.R.)$ is expected to depend on the slope of the Frank-Starling curves $F_r(P_1)$ and $F_l(P_3)$ [1], [4] according to the observation of the seven-state model.

7.1.3 What other non-oxygen saturation covariance forms need to be observed?

The covariance between oxygen saturation and various pressures as well as the covariance between oxygen saturation and heart rate were observed. The non-oxygen saturation covariance would be a good direction for further research.

7.2 Factor Analysis

The basic assumption of the factor analysis is that each factor is a linear combination of the observed variables. Whether or not there exists a function F and some $p \leq 12$ such that

$$F : (f_1, \cdots, f_p) \longmapsto (A_1, \cdots, A_{12}), \text{ for } 1 \leq i \leq p,$$

where f_i are factors, and A_i are the observed variables. The function F, expected to be linear, is associated with matrix U.

Are there more appropriate combinations, for example the linear combination of nonlinear terms? For example, different normalizations may lead to different results. Are there other normalizations except dividing by the maximun of each variable, for example dividing by Euclidean norms, and excluding outliners?

7.3 Seven-state Model

7.3.1 Model Improvement

The variables P_2, P_3, Q, H can be solved from equation 4.1-4.4. by giving fixed control terms P_1, P_4 in the seven-state model. This implies that O_c is independent of H. Some experiments show that the heart rate increases with the oxygen consumption rate [2]. How can this apparent inconsistency be solved?

One suggestion is to make the resistances R_b and R_l depend on O_c. Some experiments show that in a number of different tissues the blood flow appears to be adjusted to the existing metabolic activity of the tissue. Furthermore, imposed changes in the perfusion pressure at constant levels of tissue metabolism, as measured by oxygen consumption rate, are met with vascular resistance changes that tend to maintain a constant blood-flow rates, which is called the myogenic mechanism. Resistances also depend on neural control [P.175 in [1] and P.268 in [11]].

Another suggestion is to introduce *compliance* of blood vessels [p.109 in [8]]. The linear relation between the pressure inside blood vessels and the volume of blood vessels is

$$V_P = V_d + CP,$$

where V_P is the volume of blood vessels when the pressure inside blood vessels are P, V_d is the volume at $P = 0$, and C is the compliance of blood vessels.

The assumption that all of the pressures, flows and oxygen saturation are constant in time in the seven-state model is too loose. It would be good to consider a dynamic model to describe the circulation [Sec. 5.11 in [8]].

7.3.2 Frank-Starling Curves

The seven-state model requires the use of the Frank-Starling curves. These curves are thought to be significant in the diagnosis and treatment of heart failure. Fig. 34-2 in [5] shows the relation between Frank-Starling curves and medicine treatment. How can the medicine treatment be included into the seven-state model? Can the measured data and this model help to approximate Frank-Starling curves?

Acknowledgment

The author hereby acknowledges the support of Minnesota Center for Industrial Mathematics (MCIM) on this project. The author thanks Tom Bennett and Teresa Ruesgen of Medtronic Inc. for offering the data, relevant information and technical help. This project proceeded under the guidance of Blaise G. Morton. The seven-state model was developed after many discussions with

him. N. V. Krylov gave some advice on the covariance analysis of the seven-state model. The author sincerely thanks them and others who ever gave help on this project.

References

[1] R. M. Berne and M. N. Levy, *Cardiovascular Physiology*, (Mosby, 1992).

[2] Alan C. Burton, *Physiology and Biophysics of the Circulation*, (1965).

[3] C. Eckart and G. Young, 1936, The approximation of one matrix by another of lower rank. *Psychometrika*,1:211-18.

[4] R. L. Fisk, *The Right Heart*, (1987).

[5] Goodman and Gilman, *The Pharmacological Basis of Therapeutics*, Ninth Ed. (1995).

[6] A. C. Guyton, *Circulatory Physiology: Cardiac Output and its Regulation*, (1963).

[7] J. R. Hampton, *Cardiovascular Disease*, (1984).

[8] F. C. Hoppensteadt and C. S. Peskin, *Mathematics in Medicine and the Life Sciences*, (Springer, 1996).

[9] Bertram G. Katzung, *Basic and Clinical Pharmacology*, (Appleton and Lange, 1992).

[10] N. V. Krylov, *Introduction to the Theory of Diffusion Processes*, (AMS, 1991).

[11] A. Noordergraaf, *Circulatory System Dynamics*, (1978).

[12] A. V. Oppenheim and R. W. Schafer *Discrete-Time Signal Processing*, (1989).

[13] R. Reyment and K. G. Joreskog, *Applied Factor Analysis in the Natural Sciences*, (1996).

[14] L. B. Rowell, *Human Cardiovascular Control*, (1993).

[15] R. F. Schmidt and G. Thews (Eds.), *Human Physiology*, (1987).

[16] P. K. Thomas, L. Shure, and J. N. Little, *Signal Processing Toolbox*, (1994).