Prediction of Cardiac Arrest Outcomes Using Digital Signal Processing of Initial Electrocardiograms

Permalink
https://escholarship.org/uc/item/6qf5f3mc

Author
Lober, William B

Publication Date
1992-04-01

License
CC BY-NC-ND 4.0
Prediction of Cardiac Arrest Outcomes
Using Digital Signal Processing of Initial Electrocardiograms

By

William B. Lober

B.S.E.E. (Tufts University) 1980

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

HEALTH AND MEDICAL SCIENCES

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA at BERKELEY

Approved:

Chair:

Date 5-19-92

Date 5-19-92
Copyright Notice

Prediction of Cardiac Arrest Outcomes
Using Digital Signal Processing of Initial Electrocardiograms

Copyright © 1992

William B. Lober
Preface

My research for this project involved two distinct stages. Before I could compare patients’ electrocardiograms and outcomes, I had to first convert the electrocardiogram data into a usable form and then characterize it using signal processing techniques. I began the project thinking that I would focus on developing predictive tools for clinical, operational, or health policy settings. Yet, I now feel that my general process for applying digital signal processing techniques to field electrocardiogram data may be more immediately useful.

I believe that this is the first time that a significant sample of pre hospital electrocardiograms have been digitally analyzed, and that the results have been compared with emergency medical system data. Since collection of the underlying data is now a part of the quality improvement process, the size of the dataset and the significance of the findings will continue to increase as more data accumulates.

Both the measurements described in this paper and the techniques to develop new measurements may be useful to other investigators. I hope that my work will make field electrocardiogram data more accessible for computer analysis, and that the Data Handling Techniques section of this thesis will be helpful as a “how-to” manual for any who might wish to carry this process further.
Acknowledgments

I am indebted to a number of people on both the UCSF and UC Berkeley campuses for their enthusiasm and their assistance with my work.

Dr. Odelia Braun, director of the Center for Prehospital Research and Training (CPRT) at UCSF, allowed me complete access to the research data which CPRT collects from the San Francisco Fire Department’s early defibrillation program. She, Greg Myers, Wes Valentine, and Doug Clark gave generously of their time to answer my questions, explain the data collection and storage system, and suggest useful avenues for the analysis of the data.

Dr. David Auslander of the Department of Mechanical Engineering at UC Berkeley graciously shared with me the technical documentation he had acquired and the software he had developed to read the Laerdahl module datafiles.

I was fortunate that Dr. Walter Freeman of Molecular and Cell Biology at UC Berkeley offered me access to the custom Igor procedures developed by students and programming staff in his laboratory. Furthermore, he introduced me to Leslie Kay, a graduate student in his lab, with whom I met regularly to compare notes as we each learned to use Igor.

Finally, Dr. Auslander served as my thesis committee chairman, and Dr. William Redfearn of Biostatistics/School of Public Health and Dr. Alan Steinbach of the Joint Medical Program served as my thesis advisors. All three patiently provided valuable suggestions, encouragement, and moral support. I am very grateful to each of them.
# Table of Contents

Preface........................................................................................................... ii  
Acknowledgments.......................................................................................... iii  
Table of Contents............................................................................................ iv  

## Introduction

1. Policy Issues and Cardiac Arrest .............................................................. 2  
2. Patient Care Issues and Cardiac Arrest .................................................. 3  
3. Physiologic Markers ................................................................................ 4  
4. Early Defibrillation and Field Electrocardiograms .................................. 6  

## Problem Statement

5. Computer Analysis of Electrocardiogram Data ........................................ 8  
6. Developing New Physiologic Predictors .................................................. 8  

## Objectives

7. Develop Data Handling Methods ............................................................. 9  
8. Develop and Assess New Measurements ............................................... 9  
9. Evaluate associations between measurements and survival .................. 9  

## Materials and Methods

10. Study Design ............................................................................................ 10  
11. San Francisco .......................................................................................... 11  
12. The Heartstart 2000 Automatic Defibrillator ........................................ 12  
13. Data Collection ........................................................................................ 13  
14. Emergency Medical System Data ............................................................ 14  
15. Measurements from Electrocardiogram Data ....................................... 17  
16. Statistical Analysis Techniques ............................................................... 18  

## Data Handling Techniques

17. Flow of Data ............................................................................................. 19  
18. Preparation of Electrocardiogram Data ................................................ 20  
19. Signal Processing of Electrocardiograms .............................................. 23  
20. Development of the Analysis Dataset .................................................... 25  

## Results

21. Digital Signal Processing Measurements .............................................. 29  
22. Repetition of Prior Peak to Peak Amplitude Results ............................ 30  
23. Univariate Comparison of Measurements with Outcome .................... 32  
24. Multivariate Scatterplots ........................................................................ 33  

## Discussion

26. Issues in Methodology ............................................................................ 37  
27. Discussion of Results .............................................................................. 40  
28. Limitations of Study ................................................................................ 50  

## Conclusions

References........................................................................................................ 52  
Appendices ...................................................................................................... 54  
1. Summary of Commercial Software Used .............................................. 57  
2. Program Listing - Module Conversion Software .................................... 59  
3. Program Listing - Igor Procedure Software .......................................... 64  
4. Program Listing - SYSTAT Command Software ................................... 72  

iv
Introduction

Sudden Cardiac Death (SCD) has been defined as “A natural death...heralded by an abrupt loss of consciousness less than one hour after the onset of acute symptoms...in an individual in whom death (is) unexpected”[1]. In 1985, 560,000 adults over age 24 died of ischemic heart disease in the United States[2]. 310,000, or about 56%, died either out of a hospital or in the emergency department. Gillum identified this figure as an estimate of the incidence of SCD and described a 19% decline from 1980 to 1985. Myerberg says that the annual risk of SCD varies from 0.1-0.2% in the general population to 25% for patients who are survivors of a previous out of hospital cardiac arrest[3], and another study found a five-fold increase in incidence of SCD for patients with cardiac failure compared to the general population[4].

Cobb identified ventricular tachycardia and fibrillation as the main causes of SCD[5], and Weaver showed that ventricular fibrillation is 30 times more common in the pre hospital care setting[6]. Encouraging outcomes for patients in Seattle presenting with ventricular fibrillation include rates of hospital admission (58%), hospital discharge (29%), and discharge without serious neurological impairment (26%)[7]. These rates represent a substantial improvement following the introduction of paramedic care to Seattle about 15 years ago, and have since remained stable.

The number of patients presenting to the Seattle Fire Department in a rhythm of ventricular fibrillation decreased 25% from 6/10,000 in 1979 to 4.5/10,000 in 1985, despite the implementation of an early defibrillation program which resulted in earlier recording of patients’ electrocardiograms[7]. Cobb attributes the overall decline to a nationwide reduction in age-adjusted mortality from coronary heart disease, which he suggests may be a function of a reduction in out of hospital deaths. Out of hospital SCD remains an important cause of mortality in the United States. In the past 20 years, western medicine has made substantial progress in treating SCD. However, its treatment is expensive in a time of increasing competition for health care resources.
1. Policy Issues and Cardiac Arrest

Our society is looking at the high cost of health care at many levels. The Congressional Office of Technology Assessment is looking at the cost and efficacy of medical procedures. Oregon is trying to respond explicitly to cost containment pressures by implementing a rationing plan, which ranks all medical treatments on a community-generated scale of "value". In the United States, access to health care has traditionally been viewed as a right and explicit rationing has met with resistance both in Oregon and elsewhere. On a local level, in 1989 Alameda County was unable to adopt a consensus approach to rationing, and continues to limit costs by the "traditional" method of limiting the availability of services[8].

Cardiac arrest is an expensive disease to treat. One study from Sweden estimated that the cost for each life saved was $14,700, which included both emergency medical system and hospital costs[9]. This study was based on one year's experience with a small early defibrillation program in a health care system very different from that of the US. One US study showed the emergency medical system costs alone to be about $2200 per life saved in an early defibrillation or paramedic based system[10]. Another study showed a cost of $8900 per year of life saved in a paramedic system, which contrasted favorably with similar costs-effectiveness figures for heart, liver, and bone marrow transplantation, and for chemotherapy for acute leukemia[11].

Treatments for cardiac arrest are subject to the same evaluation as other medical interventions. Recent articles have found survival rates close to zero for cardiac arrest patients who arrest outside the hospital and are not successfully resuscitated before arriving at the emergency department[12][13]. Gray concluded that discontinuing these unsuccessful resuscitations would have saved about $1650 per patient in hospital costs.

Governmental, business, legal, and ethical interests are joining medicine in examining how health care resources are allocated. Implementing a policy based on the location and lack of success of a resuscitation effort may seem to offer attractive cost benefits.
However, I question these criteria as the best predictors of futility. An indicator of the underlying physiologic state of the patient's heart might better predict futility, and would also apply to patients who arrest in the hospital.

2. Patient Care Issues and Cardiac Arrest

In addition to the social issues, there are clinical reasons to explore physiology as a predictor of success or failure. There are many different interventions which compete for the attention of researchers and policy makers. Citizen CPR, dispatcher directed CPR, early defibrillation, several different emergency medical system designs, and various cardiac and neurological drug interventions all have their advocates. Survival is a result of many successful interventions along the path from initial resuscitation to hospital discharge. Understanding this path by exploring some of the physiological midpoints may help place these individual interventions in context.

While therapies for cardiac arrest continue to evolve, Cobb observes that since 1975 the rate of both hospital admissions (58%) and discharges (29%) following a resuscitation from ventricular fibrillation arrest has remained constant in Seattle[7]. He speculates that perhaps there is a saturation factor which applies to urban areas, that perhaps the rate is as good as it can get. He also observes that the overall rates have not improved despite the fire department's implementation of an early defibrillation program around 1984. Yet, Weaver, using Seattle data from 25 months beginning in May of 1984, found that 30% of the patients treated in the early defibrillator program survived to hospital discharge[6]. He contrasted this with a 19% discharge rate for those treated by paramedic defibrillation, who had about a 1 1/2 minute delay in response time. A physiologic indicator might help explain why a successful intervention like early defibrillation has had little overall effect.

In addition to assessing interventions, there may be an important role for physiologic measurements in direct patient care. Perhaps there are different sub populations of ventricular fibrillation patients for whom different interventions will be most effective.
3. Physiologic Markers

3.1. Amplitude

Peak to peak amplitude of human data from the emergency medical system has been examined several times as a predictor of survival. Weaver looked at 394 patients in ventricular fibrillation.[14]. He found no correlation between two categories of amplitude and clinical history, antiarrhythmic drugs, or bystander CPR. However, he did find an association between amplitude and several factors, which include cardiac arrest that was witnessed, shorter times to CPR and first defibrillation, and hospital admission and discharge.

Several years later, Weaver again looked at amplitude as a categorical variable, this time in conjunction with the early defibrillation program[6]. Of 276 patients discovered in ventricular fibrillation and evaluated with an automatic defibrillator, significant determinants of survival were witnessed collapse, coarse fibrillation (>0.2 mV), and receiving a shock from the automatic defibrillator. Not significant were younger patient age or paramedic response time.

Callaham has looked at amplitude as a numeric variable, and also examined episodes of refibrillation by the patient[15]. He found that, of 265 patients, factors associated with hospital admission were amplitude (mean 1.0 mV vs. 0.8 mV, p=0.03), witnessed arrest, stable conversion of the patient’s rhythm, and a blood pressure with the first conversion. Not strongly associated (p>0.05) with admission were time to defibrillation (mean of the two groups were similar), sex, bystander CPR, and refibrillation. For hospital discharge, he found the following differences: bystander CPR and percent change in amplitude during refibrillations were both associated with survival, while witnessed arrest was not.

Peak to peak amplitude measurements are associated in all three studies with survival. However, despite this association, there are some problems with measurement of any amplitude based quantity. Ewy points out that amplitude depends on variations in quality, shape, and location of electrode contact; skin condition and chest shape; and the direction of
the cardiac vector during fibrillation[16]. Therefore, it would be useful to find a
classification equally strongly associated with favorable outcomes which was not subject to
these interactions.

3.2. Frequency

Several investigators have examined frequency characteristics of ventricular
fibrillation, primarily in the context of detection algorithms for hospital and automatic
defibrillator instrumentation. Forster looked at field data from 206 patients in Seattle[17].
He examined ratios of frequencies which were grouped into “low”, “VF”, and “high”
bands. He did not look at outcome data, but he was able to identify ventricular fibrillation
with a positive predictive value of 0.99 by examining the power ratio of VF band (3.5 - 8
Hz) to low band frequencies.

Other investigators have found similar peak frequency results. Using 61 recordings
from 49 patients, Nolle found peak frequencies of 2.7-4.5 Hz[18]. He also found that the
distribution of the frequencies was bimodal, with a second and weaker concentration of
energy at 9-10 Hz. Murray examined 67 recordings from 61 patients, and observed a peak
frequency of 5 Hz, with a range from 2-7 Hz[19]. Ten percent of the cases had two
closely spaced peaks in their power spectrum distributions, and 8 of 10 episodes which
followed ventricular tachycardia had peak frequencies of only 2-3.5 Hz.

Herbschleb noted a peak frequency of 6 Hz. for ventricular fibrillation, about double
the 3 Hz. which he found for ventricular tachycardia[20]. Widman examined
electrocardiograms which had been identified as “homogeneous for noise and signal
content” by a cardiologist, and observed a peak frequency of 4.7±1.5 Hz[21]. He also
noted a peak frequency of 2.7 Hz. for monomorphic ventricular tachycardia, and noted
overall that there was considerable variation by individuals.

Aubert found a narrow auto correlation spectrum during ventricular fibrillation in
dogs, for both the electrocardiogram and the left ventricular pressure[22]. This indicated
that the electromechanical activity during fibrillation had a strong repeating component.
Herbschleb had looked at both dogs and people in an earlier study and found that, with a peak frequency of 12 Hz., the energy in a fibrillating dog heart was at frequencies about twice that of a human heart[23]. Arredondo showed by regression analysis that peak frequency was related to both the refractory period and the conduction time of the myocardium[24]. However, he also showed that the peak frequency was maintained for long periods of time (30 min.) in “stable” conditions, suggesting that refractory period, conduction times, and other physiological performance characteristics may not be strictly related to the length of ventricular fibrillation.

Brown and Dzwończyk found that the median frequency of the electrocardiogram could be used to estimate the time from the beginning of ventricular fibrillation in anesthetized swine[25][26]. Like Arredondo, Brown and Dzwończyk suggest that frequency characteristics may indicate the underlying cardiac physiology. Brown observes that, in a patient who has been hypoxemic for long enough to deplete their myocardium of high energy phosphate compounds, defibrillation causes an increased rate of post shock pulseless rhythms. He postulates that other therapies may be more appropriate initially than defibrillation. Since myocardial hypoxia can not be assessed in field studies, he proposes that the median frequency of ventricular fibrillation might be a proxy for time from arrest and consequent myocardial viability.

4. Early Defibrillation and Field Electrocardiograms

Early defibrillation has been reported by Weaver, among many others, to be an especially important intervention in promoting favorable outcomes from cardiac arrest[6]. However, automatic defibrillators have another benefit which has not yet been as widely used.

Unlike most manual defibrillators, the Laerdahl\(^1\) and Physio-Control\(^2\) automatic devices make both voice and electrocardiogram records for quality improvement auditing. These records can also be used as a source of research data. Dzwończyk mentions the

\(^1\) Laerdahl Corporation, 1 Labriola Ct., Armonk, NY 10504
\(^2\) Physio-Control Corporation, Redmond, WA.
difficulty of translating his results from laboratory swine to clinical medicine due to the
difficulty of obtaining electrocardiogram recordings of patients in cardiac arrest. He
suggests using recordings from patients who are in “no-code” status, and from portable
heart monitors. The data from automatic defibrillators is well suited for this purpose.

However, it can be difficult to translate this data into a usable form. Depending on
the model of defibrillator, electrocardiograms are recorded on either a cassette tape, or both
a tape and a digital storage module. If the tape data is used, then it is subject to degradation
by playback and digitizing. Also, this task requires special instrumentation and technical
sophistication on the part of the operator. If, however, the digital module data is used, then
that data must be decoded and converted to a format which other software can accept. The
commercially available Laerdahl software for their defibrillator\(^3\) will create a printout of the
module data, but it does not create an easily usable data file which can be exported to
another application.

The result of automatic quality improvement recordings is that electrocardiogram data
on cardiac arrest patients is more uniformly available. Also, times are recorded along with
the data, which is important for research purposes. However, because of the difficulty in
converting the data to a computer accessible form, the simpler measurements have been
most often studied. The characteristic of the electrocardiogram which has been most
widely reported for field data is peak to peak amplitude, which can be manually measured
from a printout. More sophisticated analysis techniques may help to detect better
physiologic markers.

\(^3\) [HeartStart Database Management Software, Laerdahl Corporation](#)
Problem Statement

A better understanding of cardiac arrest would help improve public health policy, better standards of clinical care, and promote good decision making in the care of individual patients. However, two obstacles to that understanding are the lack of well understood electrophysiologic predictors, and the difficulty in using existing field data to closely examine the electrocardiographic characteristics of different patients.

5. Computer Analysis of Electrocardiogram Data

There is a need for a general method to analyze electronically readable electrocardiogram data, both for study and for comparison with other data from the emergency medical system. There are two methods for converting the electrocardiogram data to a universal interchange format. The first method is to replay the tape recordings of the call and digitize the tape signal. This approach requires an analog to digital data acquisition system, and the engineering expertise to operate it. The second method, available with defibrillators which record the data in digital storage modules, requires custom software to convert the module data into a more general format. Currently, this method also requires technical expertise to develop and run the software. Though the underlying hardware exists, both methods remain experimental. Neither lets the user easily load field electrocardiogram data into a computer in a general format.

6. Developing New Physiologic Predictors

Because of potential problems comparing amplitudes from different clinical situations, an ideal predictor would be unitless, or at least not dependent on the voltage of the signal. This type of predictor might come from the frequency domain, from the ratios of different amplitude characteristics, from some other electrophysiologic measurement, or from a combination of several of these sources. While useful directions can be established in the laboratory, the most important assessment of these predictors must come from comparison with human field data.
Objectives

In this project, I addressed the problems of poor computer access to electrocardiogram data and the need for new physiologic predictors in three ways. I developed a method for handling digital data, defined and validated new measurements, and assessed these measurements as predictors of outcome.

7. Develop Data Handling Methods

I developed a method to reconstruct original electrocardiogram data from Laerdahl digital storage modules and produce a time series vector in ASCII format. I have also explored several specific graphical and signal processing analysis software packages on the Apple Macintosh. Using one of these, I have developed a method to isolate and analyze data from the electrocardiogram of the presenting rhythm.

8. Develop and Assess New Measurements

8.1. Amplitude Domain

I computed several amplitude domain measurements, including peak to peak amplitude, signal range, and root mean squared (RMS) voltage. I then compared these measurements with peak to peak amplitude as measured on the same data using traditional manual techniques to validate the digital measurement algorithms.

8.2. Frequency Domain

I also computed peak frequency as a measurement of the frequency domain of the electrocardiogram, and compared the peak frequency with the peak to peak amplitude.

9. Evaluate associations between measurements and survival.

Finally, I compared my measurements with survival, as indicated by hospital discharge, to assess both the univariate and bivariate relationships between these variables and the outcome of a cardiac arrest.
Materials and Methods

The subjects of this study were selected from a set of patients who had been previously studied [15]. The authors of that study compared the amplitudes of ventricular fibrillation, the outcomes of cardiac arrest, and the patients’ responses to treatment by early defibrillation. They measured amplitude by hand, using a ruler and printouts of the electrocardiograms. I chose to use the identical study group both because I was using a new technique for measuring amplitude and because I was developing new measurements of the electrocardiograms. I felt that direct comparison with Callaham’s results would help assess my methods and findings. All data used in this analysis were collected by staff of the Center for Prehospital Research and Training (CPRT) at the University of California, San Francisco.

10. Study Design

This is a retrospective cohort study, based on a complete sample consisting of all consecutive cardiac arrest patients in the city of San Francisco during the period February 1989 through December 1990 for whom recorded electrocardiogram data was available.

For Callaham’s study, the patient selection criteria were:

1) Age greater than 18 years old.
2) Etiology was non-traumatic and patient was normothermic.
3) Arrest occurred between 2/1/89 and 12/31/90.
4) First responder EMT-D care was provided by San Francisco Fire Department.
5) Presenting rhythm was ventricular fibrillation, as determined by human interpretation of the electrocardiogram recorded by an automatic defibrillator.

There were 265 patients in Callaham’s study. I enrolled 120 patients for whom I could obtain digitally recorded electrocardiogram data. This subgroup is a random sample of Callaham’s population. There were four other screening criteria which cases had to meet to remain enrolled in my study.
6) I rejected 14 cases due to either conflicts in the naming of module data files or multiple calls recorded on a single module. This left 106 cases.

7) In 4 cases, I did not know the patient’s outcome. This left 102 cases.

8) I dropped 9 cases because of gross disagreements between Callaham’s and my measurements of the peak to peak amplitude, as I felt that a large mismatch of these measurements indicated an incorrect identification of the presenting rhythm for the case. This left 93 cases.

9) The algorithm which I used to compute the power spectrum distribution required a presenting rhythm of at least 5.12 seconds in length. Five cases lacked data of sufficient length, leaving me with 88 cases.

Ninety three cases, or 78% of my original enrollment could be used in comparisons based on amplitude characteristics, while only 88 cases, or 73%, could be used where frequency measurements were indicated.

11. San Francisco

11.1. Demographics

"...San Francisco is a hilly urban environment of 49 square miles with a population of 763,800. The residents have a median age of 35.8 years. Forty percent of the population is aged 25 to 44, 4.5% are 60 to 64, 8% are 65 to 75, and 7% are over 75. The city’s ethnic mix is 54% white (including 14% Hispanic), 29% Asian, 11% black, and 6% other. (From US. Census Data)..." [15]

11.2. Emergency Medical System

The San Francisco Fire Department (SFFD) is the first responder for medical emergencies in the city and county of San Francisco. Fire vehicles responded to 1166 possible cardiac arrest calls during the study period, 880 of which resulted in resuscitation being initiated[15].

The city has a single tier emergency medical system with a triage medical dispatch center staffed by paramedics. The first responders are of EMT-D level firefighters. Advanced life support is provided by a municipal paramedic system run by the Department of Public Health, with backup response by private ambulance company paramedics. Fire
department response times average four minutes, and paramedic response averages 10 minutes.

11.3. Early Defibrillation Program

All data used in this study came from the early defibrillation program of the San Francisco Fire Department. In 1990, the SFFD operated 41 engine and 2 rescue companies as medical first responders, all of which were equipped with a Laerdahl Heartstart 2000 automatic defibrillator.

12. The Heartstart 2000 Automatic Defibrillator

From the time it is switched on, the Heartstart 2000 makes a continuous two track cassette tape recording for quality improvement auditing. The first track is an ambient audio recording, which captures the voices of the firefighters and paramedics. The second track is an electrocardiogram recording. The defibrillator also records operating information and select periods of electrocardiogram data in a solid state digital data module.

I used the data from the solid state module rather than that from the cassette tape because of both the inaccuracies inherent in reproducing the electrocardiogram signal from the tapes and the difficulty in subsequently digitizing the signal.

12.1. Characteristics of Recorded Signals

The defibrillator has a high-impedance, AC-coupled front end. The surface electrocardiogram is therefore recorded with a minimum of disturbance from the instrumentation, and with no DC component. The signal is then amplified, filtered, given a DC offset to center it in the 5 millivolt recording range, and then digitized. The front end analog low pass filter has a roll-off frequency of 20 Hertz.

The electrocardiogram signal is sampled at a 100 hertz rate. This allows theoretical capture of signals up to 50 hertz, which is well beyond the cutoff frequency of the filter.

---

4 personal communication, Doug Clark, CPRT, University of California, San Francisco, CA 94143.
5 personal communication, Doug Clark.
6 Laerdahl Corporation.
7 personal communication, Jim Angel, R&D Engineer, Laerdahl Manufacturing Corporation, 9440 SW Tualatin-Sherwood Road, Tualatin, OR 97062.
The signal is digitized to eight bits, which allows a 19.06 µV resolution, and is recorded in three second segments.

The digital data module contains a 16 Kbyte electrically erasable programmable read only memory (EEPROM). The module can store 50 three second segments, or 2 1/2 minutes worth of digital electrocardiogram data. The defibrillator determines which segments to keep based on a priority system that gives extra weight to the presenting rhythm and to segments before and after a shock.

13. Data Collection

CPRT provides the medical supervision for the SFFD’s early defibrillation program. The collection of data for research purposes is integrated into the quality improvement process. Information about each call, such as demographic factors, interventions, and outcome, is gathered from several sources. These include the Fire Department, the Department of Public Health Paramedic Division, and the receiving hospital. The CPRT staff also collect the digitally recorded electrocardiogram data for each call.

The call review process is conducted by one of several reviewers at CPRT. Most of the calls used in this study were reviewed by a single registered nurse (W.V.)\(^8\). Some calls were reviewed by the CPRT medical director (O.B.), or by a medical student (C.C.). The reviewer employs a Laerdahl Medical Control Unit (MCU)\(^9\). The MCU uses the data module and the tape recorded by the automatic defibrillator to produce a printout which includes both the electrocardiogram and a summary of the defibrillator’s performance.

The reviewer examines the printout while critiquing the audio tape of the call. The primary purpose of the review is the generation of quality improvement letters to the firefighters. These letters provide feedback about the eventual disposition of the patient,

\(^8\) Initials for CPRT staff are:
O.B. Odelia Braun, MD, Medical Director
W.V. Wesley Valentine, RN.
D.C. Doug Clark, Research Associate
G.M. Greg Myers, Research Associate
C.C. Claudia Condie, UCSF medical student

Center for Prehospital Research and Training, University of California, San Francisco, CA 94143,

\(^9\) Heartstart 2000 Medical Control Unit, Laerdahl Corporation.
and highlight both good and bad aspects of the call. In addition to performing the quality improvement task, the person reviewing the call also collects research data. These data include interpretations of the cardiac rhythms, hand measured signal amplitudes from the electrocardiogram printout, and times of certain events, which are estimated using the audio tape and a stopwatch.

14. Emergency Medical System Data

Emergency medical system data is collected for each cardiac arrest to which the SFFD responds. These data include the circumstances of the arrest, an estimate of the time of arrest, any basic or advanced cardiac life support interventions (BCLS/ACLS), the patient's response to these interventions, the times of these interventions, and several measures of outcome. These data, which I will refer to as the EMS database, are maintained in a computer database in a combination of numeric, categorical, and comment fields. The database also includes quality improvement information from the tape review and follow-up information from the receiving hospital's chart. The data in the EMS database is organized into a number of logical "pages", one for each group of similar variables. These groups are illustrated in figure 14a.

![Diagram of EMS database pages]

**Figure 14a** - The EMS database includes information on circumstances of arrest, various response times and intervention times, specific interventions and problems, and several measures of patient response and outcome. It also includes quality improvement and follow-up information.
The page most relevant to this study is the information on the "Run Sheet" page, which is detailed in figure 14b. The entire EMS database is maintained on an Apple Macintosh IIcx using Claris FileMaker Pro10.

14.1. Demographic Data

Demographic information is obtained from the Fire Department’s run record. This includes the patient’s age, sex, and date of birth. It also includes the name of the patient’s private physician and the hospital at which that physician practices. It does not include race or socioeconomic status. Identifying information has been removed from the cases selected for analysis.

---

10 Version 1.0v1. Claris Corporation, PO Box 58168, Santa Clara, CA 95052.
14.2. Interventions

Data on interventions also comes from the Fire Department’s run record. These data include whether the cardiac arrest was witnessed and whether bystander CPR was initiated prior to the arrival of the first responders. The data also include the first responder interventions: was the patient given oxygen, was an oral airway established, was ventilation provided via bag valve mask, was the patient’s airway suctioned, how many shocks did the defibrillator deliver, and did the paramedics initiate advanced cardiac life support on their arrival? The unit numbers of all firefighters and paramedics involved in the call are recorded.

14.3. Times

Various times for the fire fighter and paramedic response are obtained from the dispatch times, the defibrillator event logs, and the run records. CPRT staff records the times while reviewing the call. The defibrillator’s internal clock can be used to reliably determine intervention times for both defibrillator and external events. The later include the arrival of the paramedics and the initiation of ACLS interventions. Since the firefighters report their arrival on scene by radio, the only uncertain time interval is the period between their arrival on scene and the time the defibrillator is switched on. For simplicity, I have followed Callaham’s lead and disregarded this time interval.

14.4. Patient Presentation, Response, and Outcome

Clinical data is also transcribed from the run sheet. These include: the patient’s initial rhythm, whether pulses were ever regained, whether a blood pressure was regained, and whether the patient vomited. Information on whether the patient reverted to fibrillation following a successful conversion is obtained from the times recorded with the defibrillator event record. The CPRT data include four outcome categories: Died in the field, Died in the ER, Admitted/Died, and Admitted/Alive.
15. **Measurements from Electrocardiogram Data**

I used Igor\(^{11}\), a signal processing software package for the Apple Macintosh, to first extract the presenting rhythm from the electrocardiogram data and then make four amplitude and frequency measurements. The peak to peak amplitude, range, and root mean squared (RMS) voltage were based on the signal’s amplitude characteristics, and the peak frequency was based on its frequency characteristics.

15.1. **Amplitude measurements**

Three of the measurements I computed were based on amplitude characteristics of the electrocardiogram. The Igor procedure code which I wrote to make these measurements is listed in Appendix 3.

To measure peak to peak amplitude, I wrote a routine which read through a time series vector of voltages, keeping track of the change between points of inflection, and saving the greatest change it encountered. The routine returned the absolute value of the largest monotonically increasing or decreasing segment of the wave.

The other two values, range and RMS voltage, were both obtained from a statistical summary routine supplied with Igor. Range is the absolute difference between the largest and smallest voltages recorded in the segment. Root mean squared (RMS) is a standard engineering characterization of a signal, which Igor also computes with a supplied routine. It represents the “average” voltage of a signal whose mean is zero. A non-zero average is obtained by treating all voltages as positive as in the following formula:

\[
\sqrt{\frac{\int [V^2]}{T}}
\]

This is essentially equivalent to the calculation of standard deviation, using integration rather than summation. For a discretely sampled signal, the RMS voltage will be very close to the sample standard deviation of the individual measurements.

\(^{11}\) Version 1.25. WaveMetrics, PO Box 2088, Lake Oswego, OR 97035. Also, see Appendix 1.
15.2. Frequency Domain Measurements

I characterized the frequency domain of the electrocardiogram signal by computing the power spectrum density vector, which consists of amplitude values for different frequencies. Then I looked for the largest value of amplitude, and determined the frequency to which it corresponded. I refer to this value as the peak frequency. It represents the frequency at which the signal contained the greatest energy.

16. Statistical Analysis Techniques

To describe the measurement variables, I used the mean, standard deviation, and coefficient of variability for each. In general I used graphic techniques suitable for exploratory analysis, followed with a statistical technique to quantify the strength of the association. To compare survivors and non-survivors for a single measurement variable, I used superimposed frequency histograms, followed by Student’s t-test to quantify the significance of the difference between the means of the two groups.

To comparing two variables, whether two measurement variables or a measurement variable and an outcome variable, I used a scatterplot which showed survivors and non-survivors with different symbols. I quantified these plots using Pearson’s correlation coefficient, with the p-value indicating global significance as tested by Bartlett’s chi-square.
Data Handling Techniques

In keeping with my goal of creating a "how-to" manual for the digital signal processing of field electrocardiogram data, I will detail the steps I took to acquire, organize, and process the electrocardiogram data, and then to create a dataset for statistical analysis by combining the electrocardiogram measurements with emergency medical system data.

17. Flow of Data

The study involved the integration and processing of several different types of data, and required four phases. I first chose the study period and identified the sample for analysis. Then I developed summary measurements of the electrocardiogram data using digital signal processing techniques. Next, using one of these computed measures, I compared my results with previous measurements to verify that I was correctly reading and decoding the electrocardiogram data files. Finally, I integrated these measurements with the emergency medical system data to investigate the possible associations between the electrocardiogram characteristics and survival. Figure 17a is a flowchart overview of the entire process.

![Flowchart](image)

figure 17a - This describes the flow of data between the different databases.
18. Preparation of Electrocardiogram Data

In order to convert the electrocardiogram to a usable form, I used a two step process. Each step required developing software for the Macintosh. First, I decoded the raw defibrillator module data to reconstruct both the electrocardiogram and a record of defibrillator events. CPRT staff (G.M.) identified the starting and ending times for the presenting rhythm. I then integrated the decoded signal with this timing data and extracted the electrocardiogram of the presenting rhythm.

18.1. Storage of Module Data

The electrocardiogram data is currently saved by CPRT in two forms. First, the printouts produced for the quality improvement review are combined with copies of the field documentation, the quality improvement letters, and a printout of all the emergency medical system data for that case. These packets of documentation are stapled together and filed by date.

Second, the data from the solid state storage module is downloaded onto an IBM PC compatible 286 laptop computer. The downloading is done using the MCU and Laerdahl’s Heartstart Database Management software. Each module’s data is saved as a separate file. These binary files are 16K bytes long, and are direct images of the defibrillator’s electrically erasable programmable read-only memories (EEPROMs). Periodically the data files are backed up on 3 1/4 inch disks for permanent storage.

18.2. Organization of Downloaded Module Data

I obtained all available binary defibrillator data files by duplicating CPRT’s backup disks. Then I copied the newest data, which had not yet been backed up, from the laptop computer. The CPRT records include the module data files for all cardiac arrest calls, not just those for which ventricular fibrillation was identified as the presenting rhythm.

I converted the files to Macintosh format using Apple File Exchange, and moved the data onto the hard disk of an Apple Macintosh IIcx. I then organized the data files in

---

12 Mitsubishi model 286 XL laptop computer.
13 Version 1.1.4. Apple Corporation, 20525 Mariani Ave, Cupertino, CA 95014.
folders corresponding to the roughly chronological order of the original PC backup disks. I refer to this system of folders as the Module database. Currently, it occupies about 30 megabytes (Mb) of disk space.

18.3. Format of Module Data Files

A single data file consists of an exact image of the 16K byte EEPROM. There is a header section, a section of recorded electrocardiogram data, and a section of event or "annotation" records, as described in figure 18a. The header section contains information on how many shocks have been given and whether or not the module data has been printed by the MCU. There are 15,000 bytes dedicated to the storage of electrocardiogram segments. Each of 50 segments occupies 300 bytes of storage. Each byte stores one sample, at a rate of 100 samples per second. Each segment, therefore, stores three seconds of data.

The event records section contains a circular buffer of 1378 bytes, which stores event records. Depending on the type of event, these records occupy between 4 and 44 bytes. Two bytes of each record contain the time at which the event occurred. The most important event types are: New Episode, Vote/Analysis Result, Defibrillation, Three Second Segment Marker, Event Marker, Analysis Sum, Real Time Clock, and Master Time Change. The storage order of the electrocardiogram data, as well as the internal operations of the defibrillator, must be reconstructed from this data.
18.4. Extraction of Electrocardiogram and Event Data

In 1990, Dr. David Auslander\textsuperscript{14}, wrote a program for the IBM PC which read the binary module data files and decoded the event data and the electrocardiogram data. I obtained the source code for his program, which was written in the language C. I then converted the program to run on the Macintosh, verified the operation of the algorithm, and enhanced the program. I produced electrocardiogram data in several formats for exploratory display and processing, and extended the reporting format for event data to include all internal events of the defibrillator.

The Macintosh program is called ReadModuleData\textsuperscript{15}. The program was compiled using Lightspeed's Think C compiler\textsuperscript{16} and its UNIX I/O emulation library. A listing of the source code is attached as Appendix 2.

18.5. Identification of Time Interval of Presenting Rhythm

Staff at CPRT (G.M.) used the record of defibrillator events from ReadModuleData in conjunction with the MCU printout to identify the starting and ending times for the presenting rhythm of each call. The goal was to retrospectively identify the interval which

\textsuperscript{14} Professor, Department of Mechanical Engineering, University of California, Berkeley, 94720.
\textsuperscript{15} © 1992 William B. Lober. Module Data processed with Version 2.9
\textsuperscript{16} Version 4.0. Symantec Corporation, 10201 Torre Avenue, Cupertino, CA 95014.
had been manually measured for Callaham’s study. In that study, the presenting rhythm interval had been limited to a period where no one was touching the patient and the defibrillator had not yet started charging its capacitors to deliver a shock.

To eliminate error between observers, all interval estimations were performed by the same CPRT staff member. The time interval was reported in absolute seconds from the time at which the machine was turned on. I chose to use absolute seconds in an effort to eliminate a 0.5 second inconsistency between the way Laerdahl’s documentation indicates that times are recorded in the solid state module, and the way the times are reported by Laerdahl’s Heartstart Database Management software.

19. Signal Processing of Electrocardiograms

I extracted the presenting rhythm electrocardiograms and computed their amplitude and frequency measurements using Igor, which is a commercially available software package for data display, data analysis, and signal processing.

19.1. Igor Signal Processing Software

Following is a paraphrased summary of a description of the Igor Signal Processing software system. The description was written by the UC Berkeley Workstation Software Consulting Group. The full description of the software is included in Appendix 1.

...Igor is an interactive environment for exploratory data analysis and graphing. It allows multiple graphic views of large datasets. Page layouts can be defined, which can include multiple graphic, text, and image elements. In addition, analysis techniques include Fourier transforms, curve fitting, integration, differentiation, and mathematical expression evaluation. Igor includes an interpreted programming language that allows automation of data importing, analysis, and printing...17

The features most important for this analysis were Igor’s data importing functions, the simple wave analysis tools, the Fourier transform as the basis for power spectrum density calculations, and the built-in programming language to automate the process.

---

17 Description from Aron Roberts, Workstation Software Support Group, Information Systems and Technology, University of California, Berkeley, CA 94720.
19.2. Making Measurements with Igor

In order to use Igor to measure the signal characteristics of an electrocardiogram, I had to transform the ASCII data into an Igor "wave", select the relevant interval of that wave, and then sequentially call four measurement routines. These routines were either internal to Igor or developed by me, using its built-in programming language.

The first routine was an internal measurement routine which provided two of the amplitude statistics, range and RMS voltage. The second was a custom routine which I wrote to sequentially traverse the wave looking for the largest peak-to-peak voltage swing within a single signal complex. The third was a modified version of a supplied Igor routine to calculate the power spectrum density of the signal. Finally, I reapplied the standard measurement routine to the output vector from the power spectrum density to measure the peak frequency.

I selected several cases to use while developing the measurements. With Igor, I created graphs of their presenting rhythm intervals, of their power spectrum densities, and compressed graphs of the each entire electrocardiogram. The graph of the entire signal provides a quick visual record of the call, indicating how long the call was, how many shocks were given, and what rhythms preceded and followed each. These graphing capabilities were very important during the initial phases of the project, when, after much trial and error with other software packages, I was finally able to explore the data interactively using Igor.

19.3. Automation of the Measurements

I wrote additional Igor programs to automate all of the above processes, first for a single electrocardiogram, and then for batches of electrocardiograms. Performing these measurements requires considerable numeric processing, and can take 5-15 minutes for each call. For logistical reasons, the final processing for this study was done on an Apple Macintosh Powerbook 140 with 4 Mb memory and a 100 Mb external hard drive. Without
a language which allowed the development of batch processing, processing the data would not have been practical.

The names of the module data files and the time intervals of their presenting rhythms were supplied as an ASCII file by CPRT staff (G.M.). To avoid transcription error, I edited this file directly to form subroutine calls with three arguments: the name of the file, and the starting and stopping times for the presenting rhythm. These subroutine calls were spliced into series of Igor shell routines, one for each batch of calls processed. Igor wrote the measurements made from each electrocardiogram into an ASCII output file. The custom software used to extend Igor is listed in Appendix 3.

20. Development of the Analysis Dataset

The final step in preparing to explore associations between characterizations of the electrocardiogram and outcomes of cardiac arrest was to build an analysis dataset. This dataset combined the measurements made using Igor with data from the EMS database. I did this by creating a FileMaker Pro database of the measurements and linking it to the EMS database through the name of the electrocardiogram data file. Then I exported a combined ASCII file of the selected demographic, intervention, time, outcome, and measurement data for each case. Finally, I used SYSTAT to read and further process this data file to build a SYSTAT dataset for statistical analysis.

20.1. Measurements Database

Processing the data with Igor produces an ASCII output file consisting of the electrocardiogram file name, and four measurements for each file. These measurements include peak to peak amplitude, range, RMS voltage, and peak frequency of the signal. These data were converted to a FileMaker Pro database, which I call the Measurements database. The schema for this database is detailed in table 20a.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>dat_file</td>
<td>the id number of the electrocardiogram data file measured</td>
</tr>
<tr>
<td>w_ptp</td>
<td>peak to peak amplitude</td>
</tr>
<tr>
<td>w_rms</td>
<td>root mean squared voltage</td>
</tr>
<tr>
<td>w_range</td>
<td>range of amplitude</td>
</tr>
<tr>
<td>wf_peak</td>
<td>peak frequency</td>
</tr>
</tbody>
</table>

Table 20a. Contents of Measurements database. All variables contain numeric values. The data are matched with the EMS database through the dat_file field.

20.2. Merging Measurements and EMS databases

The name of the electrocardiogram file was present in both the EMS database and the Measurements database. I defined new fields in the EMS database which were “look-up fields”, indexed by the electrocardiogram file name. The values for these fields were obtained when this newly expanded EMS database was updated by copying the corresponding values from the Measurements database.

20.3. SYSTAT Dataset

Next, I created an ASCII output file from the expanded EMS database. This file contained all of the system data and measurements which I planned to analyze. The identifying and numeric variables included were: data file identification number, maximum amplitude measured for Callaham’s study, the four measurements (peak to peak amplitude, range, RMS voltage, and peak frequency), and patient’s age. Categorical variables included outcome, patient’s sex, and four “yes/no” intervention or response variables: witnessed arrest, bystander CPR, successful conversion of cardiac rhythm, and pulse/blood pressure regained. Finally, I included response times to the scene for both the fire department and the paramedics.

I then used the ASCII import feature of SYSTAT to create the analysis dataset. The final two steps used SYSTAT’s built-in programming language to recode certain data fields and to screen for gross errors. The data screening was accomplished by plotting the
Maxamp (adjusted to millivolts) against the peak to peak amplitude measured on the same time interval and rejecting those points which did not grossly agree, as determined visually.

I recoded the data in three ways. I scaled the hand measured amplitude value to reflect millivolts. Also, I recoded the outcome from a four valued categorical variable into a Died/Lived two valued variable. Finally, I performed log transformations on each of the measurement variables to normalize their distributions. These steps resulted in the addition of several variable fields to the analysis dataset. The complete list of variables included in the dataset is contained in table 20b. The code for the recoding and screening is included with the SYSTAT code listed in Appendix 4.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>dat_file</td>
<td>the id number of the electrocardiogram data file measured</td>
</tr>
<tr>
<td>maxamp</td>
<td>the hand measured amplitude from Callaham's study</td>
</tr>
<tr>
<td>maxamp_s</td>
<td>maxamp scaled to give a result in millivolts.</td>
</tr>
<tr>
<td>maxamp_l</td>
<td>maxamp_s transformed by ( \log_{10} )</td>
</tr>
<tr>
<td>outcome</td>
<td>Categorical variable with a range 0-4: Died/Unknown, Died in Field, Died in ER, Admit/Died, or Admit/Alive</td>
</tr>
<tr>
<td>outcome_2</td>
<td>Recoded to Died vs. Admit/Alive</td>
</tr>
<tr>
<td>sex</td>
<td></td>
</tr>
<tr>
<td>age</td>
<td></td>
</tr>
<tr>
<td>wtp</td>
<td>wptp transformed by ( \log_{10} )</td>
</tr>
<tr>
<td>w rms</td>
<td>w rms transformed by ( \log_{10} )</td>
</tr>
<tr>
<td>w ptp</td>
<td>w ptp transformed by ( \log_{10} )</td>
</tr>
<tr>
<td>w rms</td>
<td>w rms transformed by ( \log_{10} )</td>
</tr>
<tr>
<td>w ptp</td>
<td>w ptp transformed by ( \log_{10} )</td>
</tr>
<tr>
<td>outcome</td>
<td></td>
</tr>
<tr>
<td>outcome_2</td>
<td></td>
</tr>
<tr>
<td>sex</td>
<td></td>
</tr>
<tr>
<td>age</td>
<td></td>
</tr>
<tr>
<td>w_tn</td>
<td>whether arrest was witnessed (Yes/No)</td>
</tr>
<tr>
<td>w_b_cpr</td>
<td>whether bystander CPR was initiated (Yes/No)</td>
</tr>
<tr>
<td>pconv</td>
<td>did patient's rhythm convert? (Yes/No)</td>
</tr>
<tr>
<td>bp_p</td>
<td>did patient regain a pulse and blood pressure? (Y/N)</td>
</tr>
<tr>
<td>fd_resp</td>
<td>fire department response time (to scent, not patient)</td>
</tr>
<tr>
<td>pm_resp</td>
<td>paramedic response time</td>
</tr>
</tbody>
</table>

Table 20b - Contents of SYSTAT analysis dataset. This includes the variables transferred from the expanded EMS database, which includes the electrocardiogram measurements, and additional values which are created by recoding the data from within SYSTAT.
Results

My results are presented in four parts. First, the digital signal processing measurements are presented, along with their descriptive statistics. Next, the peak-to-peak amplitude is compared with manual measurements made for Callaham's study to assess the validity of the digital measurement techniques. Finally, each measurement variable is compared with outcome, first in univariate and then in graphic multivariate comparisons.

2.1. Digital Signal Processing Measurements

The Igor software can produce both a one page graphic printout for each call and a summary disk output file containing the measurements for all calls processed.

![Graphs showing amplitude over time and peak to peak values](image)

**Figure 21a** - This is a graph printout for a representative electrocardiogram. It contains a compressed overview of the call, a printout of the presenting rhythm in context, the isolated presenting rhythm with amplitude measurements, and the power spectrum density with the peak frequency.

An example of the graphic printout is shown in figure 21a. The printout includes a compressed graph of the electrocardiogram for the entire call, which depicts both the valid
data segments and the intervals between them when the storage space was reused to allow the capture of a later segment. The printout also shows two expanded views of the presenting rhythm, one which shows the rhythm in context of the surrounding electrocardiograms, and one which shows just the extracted segment on an expanded vertical axis. The peak to peak amplitude is printed on this graph, and the range and RMS voltage are printed below it. Finally, the peak frequency is printed on the graph of the power spectrum density.

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>Mean</th>
<th>S.D.</th>
<th>C.V.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak to Peak</td>
<td>93</td>
<td>0.914 mV</td>
<td>0.496</td>
<td>54%</td>
</tr>
<tr>
<td>Range</td>
<td>93</td>
<td>0.986 mV</td>
<td>0.514</td>
<td>52%</td>
</tr>
<tr>
<td>RMS Voltage</td>
<td>93</td>
<td>0.185 mV</td>
<td>0.098</td>
<td>52%</td>
</tr>
<tr>
<td>Peak Frequency</td>
<td>88</td>
<td>3.666 Hz</td>
<td>1.24</td>
<td>34%</td>
</tr>
<tr>
<td>Hand Measured Peak to Peak</td>
<td>93</td>
<td>0.926 mV</td>
<td>0.496</td>
<td>54%</td>
</tr>
</tbody>
</table>

Table 21b - Descriptive statistics for each measurement variable. Note that there are only 88 cases which had a presenting rhythm long enough to permit frequency analysis. (n - number of cases, S.D. - standard deviation, C.V. - coefficient of variation)

Descriptive statistics for each of the four measurement variables, summarized across the entire dataset, are presented in table 21b. Also presented are summary statistics for the hand measured peak to peak amplitudes from Callaham’s data for the same set of cases. Peak frequency could not be calculated for five of the presenting rhythms because the time interval specified for each was less than the 5.12 seconds required by the power spectrum density algorithm.

22. Repetition of Prior Peak to Peak Amplitude Results

Of the four measurements, only the peak to peak amplitude and the range can be measured without a computer or other instrumentation. Of these two, only peak to peak amplitude has been used in prior research at CPRT. The CPRT staff makes a manual peak to peak amplitude measurement as part of the quality improvement review. This
measurement is made by visually inspecting the presenting rhythm for the largest peak to peak distance within a single complex and measuring the distance with a ruler. The manually measured peak to peak amplitudes were already recorded in the EMS database as part of the work for Callaham's paper.

**Manual vs. DSP Measurements**

![Diagram](image)

**Figure 22a** - This plot shows hand measured peak to peak amplitude (MAXAMP) plotted versus digitally measured amplitude (W_PTP). The ellipse excludes the points which were removed from the dataset. Solid and hollow symbols represent patients who lived and died, respectively.

Figure 22a shows the hand measured peak to peak values (MAXAMP) plotted against the digitally measured values (W_PTP). Most of the values were in very close agreement visually. However, there were a few points for which the hand and computer measured values disagreed. Since it was important to ensure that the same intervals of data were being measured, I removed these points from the dataset, and asked CPRT staff (G.M.) to recheck the identification of the time intervals for these presenting rhythms. Once the data
(points outside of the ellipse were rejected, there was very close agreement of these variables. The Pearson correlation coefficient is: \( r = 0.986 \) (\( p < .000005 \)).

23. **Univariate Comparison of Measurements with Outcome**

I compared both direct measurement variables and \( \log_{10} \) transformed variables against outcome. Outcome was treated as a dichotomous variable (Died vs. Lived), with “Lived” defined as hospital discharge. The direct variables used were the four computer measured values and the hand measured amplitude. The \( \log_{10} \) transformed values of peak to peak amplitude, RMS voltage, and the hand measured amplitude were also used.

![Comparison of Distributions of peak to peak amplitude for patients who Died vs Lived](image)

**figure 23a** - This plot compares the frequencies (histogram bars) of the peak to peak amplitudes for each of the groups. The curves represent scaled distributions for the two groups.

Figure 23a shows the distribution of peak to peak amplitude among the patients who died versus those who lived. The histogram bars illustrate that the sample distribution of survivors is completely contained within the sample distribution of non-survivors for this variable. The distributions overlap similarly for the other measurements. However, the means of the two groups may still be distinguished from each other using Student’s t-test. Table 23b presents the means and the significance of their differences.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Died</th>
<th>Lived</th>
<th>$t$, $df$</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (mV)</td>
<td>n</td>
<td>Mean (mV)</td>
<td>n</td>
</tr>
<tr>
<td>Peak to Peak</td>
<td>0.875</td>
<td>84</td>
<td>1.285</td>
<td>9</td>
</tr>
<tr>
<td>Range</td>
<td>0.940</td>
<td>84</td>
<td>1.414</td>
<td>9</td>
</tr>
<tr>
<td>RMS Voltage</td>
<td>0.177</td>
<td>84</td>
<td>0.260</td>
<td>9</td>
</tr>
<tr>
<td>Peak Frequency</td>
<td>3.600</td>
<td>79</td>
<td>4.25</td>
<td>9</td>
</tr>
<tr>
<td>Hand Measured Peak to Peak</td>
<td>0.884</td>
<td>84</td>
<td>1.267</td>
<td>9</td>
</tr>
</tbody>
</table>

Measurements transformed by Log$_{10}$ to ensure normality

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak to Peak</td>
<td>-0.128</td>
<td>84</td>
<td>0.100</td>
<td>9</td>
</tr>
<tr>
<td>RMS voltage</td>
<td>-0.082</td>
<td>84</td>
<td>-0.601</td>
<td>9</td>
</tr>
<tr>
<td>Hand Measured Peak to Peak</td>
<td>-0.123</td>
<td>84</td>
<td>0.092</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 23b - Results of Student’s t-test for the direct and log-transformed measurements (n - number of cases, $t$ - t statistic, $df$ - degrees of freedom, p value - significance). $^*$ indicates p value based on pooled variances, otherwise separate variances were used.

Because of the small sample size and the unequal size of the groups, the assumptions that the two groups be normally distributed and have equal variances are especially important for Student’s t-test to be accurate[28]. I transformed the peak to peak amplitude and RMS voltage variables by log$_{10}$ in order to more closely fit the data to a normal distribution. The distribution of the frequency variable was normal without transformation. Because the unequal group size, I compared the group variances using the F test [28]. Equal variances were rejected (at $\alpha = 0.05$) for all variables except RMS voltage and peak frequency. Therefore, pooled variances were used for the significance tests of these two variables, and separate variances were used for tests of peak to peak frequency, range, hand measured amplitude, and all log$_{10}$ transformed variables.

24. Multivariate Scatterplots

The difference between the means was most significant for the log$_{10}$ transformed peak to peak amplitude. I explored this variable graphically by plotting it against variables representing each of the other three measurements. The resulting graphs reveal both the
bivariate relationships between pairs of measurements and the clustering of non-survivors and survivors. The hollow symbols represent the patients who died and the solid symbols represent the patients who lived. “PTP” is used as an abbreviation for peak to peak amplitude.

![Log Range vs Log PTP](image)

Figure 24a - This plot compares the log\_10 of peak to peak amplitude with the log\_10 of the range.

Figure 24a compares the log\_10 of the peak to peak amplitude and the log\_10 of the range. The two variables are highly correlated, with a Pearson correlation coefficient of: \( r = 0.986 \ ( p < .000005 ) \). The patients who lived are clustered among the higher values of both variables, but the cluster is within the region of values for non-survivors. The use of these two variables together does not provide any distinct visual separation of the two groups.
figure 24b - This plot compares the log$_{10}$ of the peak to peak amplitude with the log$_{10}$ of the RMS voltage.

Figure 24b compares the log$_{10}$ of the peak to peak amplitude with the log$_{10}$ of the RMS voltage. These variables are not as highly correlated as those of figure 24a, but still indicate a very strong relationship. The Pearson correlation coefficient is: $r = 0.971$ (p < .000005). Again, the survivors are clustered in the higher values of both variables, and again there is no obvious visual delineation between the survivors and non-survivors. However, taken together, these two variables taken together show a tighter clustering of survivors around the lower values of their range.
Figure 24c compares the log_{10} of the peak to peak amplitude with the peak frequency. Though the peak frequency is a numeric variable, the power spectrum algorithm yields discrete frequency results, giving a categorical appearance to the data. The correlation between the variables is weak: r = 0.200 (p = 0.061). All of the survivors appear in the higher peak to peak amplitudes, and they tend to be clustered in the middle frequencies. Again there is no discrete visual separation of the two groups, but again there is a tendency for the survivors to be found in a cluster. All three graphic comparisons demonstrate regions, for this sample, where survivors were not found.
Discussion

26. Issues in Methodology

26.1. Study Design

In choosing to adopt Callaham’s study population, I made an implicit decision to begin with his study design. While replicating his study sample enabled me to perform a robust assessment of my measurement techniques, I feel that it remains important to explore the assumptions involved in choosing that sample.

Callaham’s study was a retrospective cohort study, based on a complete sample consisting of all consecutive non-traumatic, normothermic, cardiac arrest patients in the city of San Francisco during the period February 1989 through December 1990. Since his sample was this entire population, rather than a stratified or random subset, population statistics can be calculated directly instead of estimated.

My study group is a random sample of Callaham’s population. My intent was to enroll his entire study population. This could not be done, as electrocardiogram data were not available for every call. However, the missing data appear to have been lost in a way which was not related to any aspect of the patient’s cardiac arrest or their outcome. In other words, while not a true random draw, the basis by which data were selected for this study was random with respect to the variables being examined. I feel that the following five criteria, which were used to select my sub-sample, are unlikely to have introduced any bias into the data.

1) I could not find electrocardiogram data for 145 of the original 265 cases. Most of the data lost were from the beginning and ending parts of the study period, although some data were lost throughout the period. I believe these missing cases reflect two types of problems in accumulating the data. First, it seems likely that there were start-up problems in establishing the data collection system. Second, there may have been additional problems once the database grew to a critical size, for instance once the data collected outgrew the laptop’s
hard disk. In either case, it is unlikely that the missing cases are biased by any characteristics of the cases themselves. There were 120 cases with electrocardiogram data available. This included none of the calls between 2/89 and 8/89, 87% of the calls between 8/89 and 5/90, and 20% of the calls from 5/90 through 12/90.

2) Fourteen cases either had CPRT-assigned names which conflicted with my software’s internal naming scheme, or they had multiple cardiac arrest calls recorded on the same module. The latter problem caused a different type of naming conflict with my software. These cases were distributed throughout the study period, and were rejected due to internal software limitations instead of case characteristics.

3) In four cases, the patient died but the EMS database did not contain information as to whether the death occurred in the field, the Emergency Department (ED), or the hospital. These cases were dropped from analysis pending verification of the outcomes, which did not happen in time to re-enroll the cases in this study.

4) Nine cases were dropped due to a gross disagreement between Callaham’s and my measurements of the peak to peak amplitude. I removed these cases from the sample so the time intervals of the presenting rhythm could be verified by CPRT staff. In all cases, the reviewer noted errors in their original estimation. However, I was unable to reprocess those cases in time to include them in this study. I believe the reviewer is likely to have made the initial errors randomly. This is supported by the fact that eight of the cases were patients who died, and one was a patient who survived, which is roughly the proportion of the overall dataset.

5) Five cases lacked the minimum of 5.12 seconds of presenting rhythm data required by the power spectrum distribution algorithm I used. There are three reasons why this might occur. First, there might be an error in identifying or
transcribing the presenting rhythm interval times. Second, in cases where an unusually large number of shocks were delivered, these segments of the electrocardiogram storage memory may be reused. Finally, the homogeneous portion of the presenting rhythm may be short because the neighboring electrocardiogram signal contains a large component of artifact and is therefore unsuitable for frequency analysis. Either way, these cases were only rejected from those comparisons which required the peak frequency measurement; they were included in the comparisons which used only amplitude measurements.

26.2. Target Population

The definition of “target population” is the group to which the results of a study may be meaningfully extrapolated[27]. By contrast, the “sampled population” is that subset of the target population which is available to be studied. For instance, if the target population is all adults suffering cardiac arrest, the sampled population might be all adult emergency department cardiac arrest patients in a specific city, during a specific period of time. The bias introduced by selection of the sampled population is as important as the sample bias in understanding how to generalize the results of a study.

In this study, the target population is all adults who suffer a non-traumatic and normothermic out of hospital cardiac arrest. In contrast, the sampled population was all cardiac arrest patients who met the following three conditions:

1) The patient was over age 18 and had a cardiac arrest in the City of San Francisco which was non-traumatic, normothermic, and occurred out of hospital. This study did not attempt to link etiology of cardiac arrest or underlying medical condition to outcomes. However, it seems reasonable to assume that there may be a link. Perhaps, for example, San Francisco has a larger proportion of young patients with opportunistic pulmonary infections who have subsequent cardiac arrests due to secondary arrest of essentially healthy hearts. Comparing age, etiology of arrest, and disease status among
different populations would be the most powerful way to show any effects of this type, but information of this type is often not uniformly collected.

2) The first responders were automatic defibrillator-equipped units of the SFFD.

3) The patient presented in an initial rhythm of ventricular fibrillation. In the EMS database for this period, there were 265 calls for which the Center staff had identified the presenting rhythm as ventricular fibrillation, from the total of 880 calls in the database. While only about 30% of the calls in this setting had a presenting rhythm of ventricular fibrillation, virtually all sudden cardiac deaths are presumed to include a progression through this rhythm [6].

26.3. Outcome Definition and the “Utstein Style”

The classification of outcomes has been a controversial topic in cardiac arrest research. A multi-national task force has released recommendations for uniform collection of cardiac arrest data for research purposes[28]. They recommend that, at a minimum, “spontaneous circulation restored” and “alive > 1 year” be recorded. Additional midpoints are “admitted to ICU/ward” and “discharged alive”. In addition, Glasgow-Pittsburgh performance scores should be obtained for patients at discharge and at one year.

The CPRT database for the study period included five outcome categories: Died/Unknown, Died in the field, Died in the ED, Admitted/Died, and Admitted/Alive. It did not include any information about neurologic status or long term follow-up.

Subsequently, the database has been expanded to include Discharged to Skilled Nursing Facility as a separate category. Also, information on cause of death in hospital and if/when the patient regained consciousness is now included, as is the hospital medical record number for further follow-up via chart review.

27. Discussion of Results

This study differs from previous examinations of ventricular fibrillation in four ways. First, I was able to apply laboratory analysis techniques to a significant sample of field data. Second, I compared the efficacy of digital and manual measurements of field data.
Third, I explored both new and previously studied measurements. Finally, using multiple measurements gave me the opportunity to examine the combined predictive value of pairs of these measurements.

Two of the measurements I chose for this paper, peak to peak amplitude [15][14][6] and peak frequency[17] have been examined previously using pre hospital data. A similar frequency measurement, the median frequency, has been explored both with laboratory[25] and field\(^\text{18}\) data. Range and RMS voltage of pre hospital data have not been examined before. Perhaps most important, this is the first time that amplitude and frequency data have been compared for a significant number of patients as a predictors of outcome.

27.1. Amplitude Measurements

Peak to peak amplitude was used by both Callaham as a numeric variable and Weaver as a categorical variable. I chose to treat it as a numeric variable in order to compare my results directly with those of Callaham.

Table 21b shows summary data for my peak to peak measurements and those of Callaham on the same set of 93 cases. His mean measurement is about 1.3% higher than mine, though the standard deviations agree exactly. The untransformed distribution of both amplitudes was slightly skewed towards zero. Figure 23a depicts this for my peak to peak measurement. Callaham’s mean result (±SD) for his entire cohort of 265 patients was 0.88 ±0.48 mV, and he reported a normal distribution for the population. The mean of my measurements was well within his 95% confidence interval for the population mean (0.82 to 0.94 mV).

Weaver in 1985 obtained an initial amplitude of 0.55 ±0.25 mV in 394 patients. In 1988, Weaver classified the initial electrocardiograms of 504 patients treated by automatic defibrillator or paramedics as fine fibrillation (≤200 mV) or coarse fibrillation (>200 mV). He found a coarse rhythm in 435 patients. My measurements were numerical rather than categorical, so direct comparison with this study is difficult. Weaver’s choice of a

\(^{18}\) Unpublished data from Chris Barton, MD, Division of Emergency Medicine, University of California, San Francisco, CA 94143.
threshold of 200 mV for the 1988 paper is three orders of magnitude higher than his 1985 measurements. I believe there must have been a printing error, and that he actually used a threshold of 200 μV, or 0.2 mV.

Range was highly correlated with peak to peak amplitude, as shown in figure 24a. This is as expected, providing that the DC offset of the signal does not change significantly during the period examined. Range may prove to be a useful and practical measurement as it is very easy to calculate. To determine range, it is only necessary to identify the maximum and minimum voltage values in the time series.

RMS voltage, range, and peak to peak amplitude have very similar coefficients of variation, indicating that the “spread” of the normal distribution which characterizes each is about the same. However, RMS voltage does not correlate quite as well with peak to peak as range does.

There are many other characteristics of the signal which are related to the amplitude, and to the distribution of different amplitudes. It is easy to imagine various statistics, such as the ratio of “high” peaks to “low” ones, the number of peaks in a period of time, or the “angularity” of the peaks. However, many of these characteristics can also be captured by examining the signal in the frequency domain.

27.2. Frequency Measurements

I choose to look at peak frequency, which is the single frequency at which the signal has the greatest energy. This is the same measurement used by several authors in both human and animal studies[17][18][23][20][19][24][21], though only Forster used data from out-of-hospital human cardiac arrests. He was interested in the power spectrum as a means of detecting ventricular fibrillation, and did not report comparable summary statistics for his sample. However, he defined the frequency band between 3.5 and 8 Hertz as the ventricular fibrillation band, based on both his and earlier work. My calculation of the sample mean (±SD) was 3.66±1.24 Hertz with a range from 0.5 Hz to 7 Hz. About 40% of my peak frequency measurements fall below Forster’s 3.5 Hz boundary. Murray
found lower peak frequencies in patients who had recently been in ventricular tachycardia. However, Forster's data came from an emergency medical system with similar early defibrillation response times, so a significant difference in duration of ventricular fibrillation seems unlikely.

An alternate way of obtaining a single number to characterize the frequency domain is to calculate a median, or centroid, frequency from the amplitude/frequency vector. The centroid is the "weighted average" frequency, weighted by the amplitudes of the individual frequency components. The difference between these two techniques of characterizing the frequencies will be greatest for signals for which the power spectrum distribution is both skewed and of relatively homogeneous amplitude, and least for signals for which the power spectrum is strongly peaked about a single frequency.

Brown used the median frequency to look at ventricular fibrillation in 11 swine. He found an interval of declining frequency from 0 to 1 minute, increasing frequency from 1 to 3 minutes, and then declining frequency from that point on. The relationships within each of these three regions were closely linear. He was interested in the time course of ventricular fibrillation and therefore did not state an overall mean estimate of his median frequencies, whose values ranged between 6 and 14 Hz. These are considerably higher than my peak frequency values. While median frequency and peak frequency should be similar for power spectrum distributions which are not markedly skewed or multi-modal, the values among different animals appear to differ. Herbschleb found that dogs had a peak frequency about twice that of humans[23]. Swine may also have a somewhat higher peak frequency. Of more interest, Brown reproduced two power spectrum distributions in his paper, one bimodally distributed and one trimodally distributed. In contrast, most of my power spectrum distributions had several minor peaks, with one clearly dominant, and a single mode distribution around the major peak.
Barton calculated a frequency measurement for 72 cases from the same series which I used. He found a mean estimate (±SD) of his measurement to be: 8.47±1.60 Hz, with a coefficient of variation of 19%. This seems to support the range of Brown’s values. However, Barton’s results were obtained from human data and Brown’s from swine.

Brown, Barton, and I used all used different software packages to compute the power spectrum distribution. Also, while the sampling rate should not make any difference, we used samples of different lengths. I do not have any information on what other windowing or pre-processing noise reduction Brown may have done. The few histograms of individual cases which he includes show a frequency spectrum with either a bimodal or trimodal distribution. This is very different from the basically “clumped” distributions which Forester and I obtained. Forester’s distributions, on the other hand, were more “spiked” than mine. Perhaps there are unresolved methodological differences between these four studies.

Again, as with the amplitude measurements, there are certainly many other useful measurements of the frequency domain. These might include the spread of the power spectrum, the distribution among specific frequencies, the number of strong frequency peaks, or the presence of specific peaks. I believe that these statistics should be pursued with consideration both of the underlying physiology which generates these wave forms and the clinical characteristics of the patients who exhibit them.

27.3. Repetition of Amplitude Results

Callaham’s paper was based on peak to peak amplitude data which was measured by hand in 1991. This variable was already in the EMS database, so it could be easily compared with the new measurements I made for this study. The availability of the hand measured peak to peak value offered a way for me to validate my technique for making digital measurements, as well as a further opportunity to uncover errors with the specification or transcription of the time intervals for the presenting rhythm.

19 Unpublished data from Chris Barton, MD.
For Callaham's study, the CPRT reviewer visually identified the largest complex in the presenting rhythm from the electrocardiogram printout, and then measured the height of the complex with a ruler. The scale of the printout is 1 mV = 10 mm. I believe the accuracy of the manual measurement technique to be 0.5 mm, or 50 µV. This is slightly less than the 38 µV precision of the difference between two digitized voltage values, each with a 19 µV precision, which is used to calculate the computer measured peak to peak.

The agreement between the two measurement systems is not perfect. The two most likely sources of errors are both attributable to the hand measurement technique. These are the error in making the hand measurement and the possibility for transcription error in recording the measurement. The component of combined error due to these sources could be estimated by measuring a single electrocardiogram several times during different quality improvement reviews, and characterizing the variation between measurements. Also, because the review process now employs more than one reviewer, there is an opportunity for "between observer" error.

I feel that the discrepancy between the hand and computer measurements was due to errors in making the hand measurements. This is not contradicted by the cases which were rejected due to gross disagreement of the peak to peak values. The CPRT reviewer (G.M.) found, when he rechecked the time intervals for those calls, that they had been incorrectly estimated. Therefore, the amplitudes were not necessarily in genuine disagreement. I have not yet compared the measurements based on the new time intervals with the hand measurements to determine if the discrepancies are resolved.

There is no between measurement or between observer variation for the digital technique. Nor is there any chance of transcription error as the data is recorded directly in the data file. However, there is a limit to the precision of the measurement which is attributable to the stepping effect of 19 µV digitizing levels.
27.4. **Univariate Comparisons**

In order to look for associations between measurement variables and outcomes, I chose to look first at survival as a simple dichotomous variable. I defined survival as being discharged from the hospital. This is problematic as neither level of function, neurological status, nor long term survival is being assessed. Also, the cause of death is not considered. Patients may have a successful cardiopulmonary resuscitation, and die at a later time from a complication such as aspiration, infection, or neurologic impairment. These deaths would be counted as unsuccessful cardiopulmonary resuscitations in this study. This dilutes the statistical effect that can be demonstrated for any factor which originally differentiates survivors from non-survivors.

A “chain of survival” analysis would strengthen the results by considering only those patients who survived to a certain stage when analyzing outcomes from the subsequent stage\(^{20}\). For instance, only those patients who had attained a pulse and blood pressure in the field would be considered when looking at predictors of survival in the ED.

Figure 23b illustrates the superimposed histograms and scaled normal distribution curves for the peak to peak amplitudes of survivors and non-survivors. I made similar plots for each of the four measurement variables, and for the \(\log_{10}\) values of peak to peak amplitude and RMS voltage. The \(\log_{10}\) plots and the peak frequency plot showed the best visual correlation between the histogram data and the normal curves.

All plots shared the important characteristic that the normal distribution for the patients who died completely encompassed the distribution for the patients who lived. Therefore, if the value for an individual patient is in the distribution of survivors, it can not be predictive of either survival or death. However, the region within the distribution of non-survivors but outside of the survival distribution could possibly be used as a predictor of death. Population estimates of the mean and standard deviation need to be developed

\(^{20}\) Suggestion from Dr. Richard Juster, Research Computing Group, Information Technology Services, University of California, San Francisco, CA 94143.
numerically for each group, in order to characterize the region of non-overlap at various associated probabilities.

The probabilities that the means of the two groups are different were developed by Student’s t-test for each of the measurement variables and log_{10} transforms. The results are summarized in table 23b. The actual p values are shown. All the pairs of means are significantly different except for peak frequency (p=0.13). The peak to peak amplitude was best at distinguishing the means, followed by the range, hand measured peak to peak amplitude, and RMS voltage. I think that the better significance of computer measured peak to peak amplitude versus hand measured is further evidence that the computer technique is more accurate.

The log_{10} transformed variables were much better at separating the means, which is consistent with my visual impression of better normality from the histograms. It is important to consider that transforming the data does not alter its inherent association with outcome, but instead simply changes the distribution of the data so it better fits the assumptions of the statistical test used. When the assumptions are better met, the statistical test becomes more accurate and associations, or the lack thereof, become more evident.

It is not surprising that peak frequency was a poor way of separating the means. Recall that Brown’s data did not show a linear relationship with time, but rather a complex curve with intervals of decreasing, increasing, and then decreasing frequency with respect to time. Others have found a correlation between time from arrest and success of defibrillation. I would not expect variables which did not change monotonically with time to separate well into groups by resuscitation success.

27.5. Bivariate Comparisons

Bivariate Scatterplots are most useful if they reveal a relationship not evident from the univariate histogram plots. The data for the survivor subgroups may be completely within the range of data for non-survivors for each of two variables, yet when plotted together, the data may reveal two clusters that can be separated by a straight line.
figure 27a - Sketch of a situation where the bivariate scatterplot would show a
delineation between groups, despite the fact that the range and distributions of
the individual groups overlap\textsuperscript{21}. The dashed line separates the two groups.

This is illustrated by the sketch in figure 27a. The equation of the dashed line could
be used to predict which group a new individual will belong to based on these two
measurements.

With four measurement variables available, there are six possible pairwise
comparisons that can be made. While I examined all six, I chose include only the three
which contained the log\textsubscript{10} of peak to peak amplitude. This was the measurement which
showed the strongest univariate relationship to survival. The first two plots, figures 24a
and 24b, were discussed earlier, along with the measurements they compare: range, and
RMS voltage. The third plot compares peak to peak amplitude with peak frequency.

\textsuperscript{21} Idea courtesy of Dr. Bill Redfearn, Department of Biostatistics, School of Public Health, University of
California, Berkeley, CA 94720.
The plot of peak to peak amplitude versus peak frequency (figure 24c) did not show a clear visual separation into two groups. However, there was evidence of clustering of the survivors in the mid-frequencies and high amplitudes. This region is also inhabited by individuals who died. I believe this scatterplot may not give significantly more information than the frequency distribution plots of peak to peak amplitude for the two groups.

This was contrary to my expectations. Brown's data shows that there is a relationship between median frequency and time. I felt that there would be a strong relationship between peak frequency and median frequency. Although the non-linearity of the median frequency makes it a poor predictor of survival when taken alone, Brown has shown that it has strong relationship with time from arrest. Similarly, others have shown that amplitude bears a relationship to time, and that time to defibrillation bears a relationship to survival. I had postulated that peak frequency would strengthen the power of amplitude to predict overall survival.

There are several possible explanations for my result. The first is that frequency may have no correlation with survival. However, Brown's paper shows a strong relationship between median frequency and time from arrest in laboratory swine. Assuming it is reasonable to extrapolate from swine to humans, and from the laboratory to the field, it seems likely that some aspect of frequency would correlate with time from arrest. Furthermore, Arredondo showed that peak frequency was linearly related to at least two underlying physiological parameters of the myocardium: refractory period and conduction time. This also supports a theoretical association between frequency and success of resuscitation.

Second, peak frequency may not have as good a correlation with time from arrest as median frequency does. The inconsistencies between the shapes of the power spectrum curves obtained by Forester, Brown, and myself is disconcerting. It would be helpful to compare methods by examining the same set of data with the variations in procedure used
by each of us. However, again Arredondo’s results suggest that peak frequency is also a
good indicator of the underlying physiologic state of the myocardium.

28. Limitations of Study

There are several factors which may limit the strength of the associations shown in
this study. The first factor is the size of the study. Among the 88 patients for whom there
was sufficient electrocardiogram data for frequency analysis, there were only 9 survivors.
In the 17 months since this study period ended there have been over 300 additional calls
added to the database. The data collection system at CPRT has now matured, the survival
rate in San Francisco has improved, more of the data recommended by the Utstein
conference are being recorded, and it is unlikely that there are as many cases for which the
electrocardiogram data is missing. All of these factors should significant power to a study
which includes the more recent data.

As discussed in section 27.4., a stepwise or “chain of survival” analysis which
considered, for example, only those patients who had regained spontaneous circulation in
the field when assessing hospital admission, or only hospital admissions when assessing
hospital discharge, could be used. This type of analysis may reduce the diluting effect of
an outcome which is the end product of many factors besides the characteristics of the
initial electrocardiogram.

Finally, the measurements themselves may be responsible for the relatively weak
associations between frequency characteristics and survival. The variations in power
spectrum curves found by several authors indicate different application of the fast Fourier
transform technique. There are choices to be made in sampling rate, in sample size, and in
windowing the data, which will change the frequency domain results. Application of
Fourier transforms to this type of data is relatively recent, and should be viewed as an
experimental rather than an established technique.

Section 27.2. postulates several other characterizations of the frequency domain. The
human brain remains much more proficient than the digital one at recognizing patterns and
trends. I believe that the best way to discover predictive frequency domain measurements would be to view a large number of power spectrum distributions and try to understand how survivors "look" different from non-survivors.
Conclusions

The original goals of this study, to create a process to digitally analyze field electrocardiogram data and to begin a search for amplitude-independent predictors of survival, have been satisfied. I have demonstrated a method of acquiring and digitally analyzing presenting rhythm electrocardiograms from pre-hospital patients in cardiac arrest. Also, I have developed several new measurements of electrocardiogram characteristics and compared them to previously published results. Finally, I have assessed these measurements, alone and in pairs, as predictors of survival.

The literature suggested that amplitude and frequency of the electrocardiogram might be associated with survival, and I developed several measurements based on these characteristics. I then obtained data on 93 cardiac arrest patients found in ventricular fibrillation, measured these values, and assessed their associations with survival using graphical and statistical techniques.

My measurements of peak to peak amplitude were consistent with those of previous studies[15][14][6]. The technique of digital measurement was highly correlated with the manual technique for the same data set, and the digital measurements showed a stronger association with survival. Range appeared to be highly correlated with peak to peak amplitude, and is easier to measure. RMS voltage was less strongly associated with both peak to peak amplitude and with survival.

My measurements of peak frequency were similar to the field measurements of Forster[17], and supported by several others[18][23][20][19][21]. However, I am concerned that the shapes of the power spectrum curves obtained by Forester, Brown[25], and myself differed in some respects. These curves are used to derive both median and peak frequency. Peak frequency was only moderately associated with survival, though taken together with peak to peak amplitude the data showed a clustering of survivors within the range of values for non-survivors. This pair of variables may have utility as a negative
predictor of survival, though the paired comparison needs to be compared statistically with peak to peak amplitude alone.

Though there are several potential pitfalls in the measurement of peak to peak amplitude, my study failed to uncover a better positive predictor of hospital discharge. The reasons for this may include the small size of the study, the difficulty in separating the many interlinked factors which comprise the “chain of survival”, methodological difficulties in measuring the frequency domain, or simply not having found the characteristic of the electrocardiogram which best discriminates between those who are resuscitated from cardiac arrest and those who are not.
References


Appendices

1. Summary of Commercial Software Used

The following product description of the Igor signal processing software package comes from Aron Roberts at Workstation Software Consulting Group, UC Berkeley.

Igor

Graphing and data analysis

Macintosh Plus or larger; 1 MB of RAM; hard disk drive; MacOS 4.1 or later.

Igor is an interactive environment for experimentation with scientific and engineering data and for the production of publication-quality graphs and page layouts. Igor allows multiple data sets of any length to be displayed in any number of graphs and tables. In graphs each curve has its own color, line size and style. Page layouts present graphs, tables, text annotation and PICTs. Fonts, type sizes and styles can be mixed in axis labels and annotation. Analysis includes Fourier transforms, curve fittings to built-in and user-defined functions, histograms, integrations, differentiations and mathematical expression evaluation. Igor includes a structured macro language that allows automation of importing, analysis, printing and other operations. Igor is extensible by C programmers using the optional External Operations (XOP) Toolkit. XOPs supplied with Igor allow serial port and GPIB communications, importation of Cricket Graph and Excel data files.

$295 retail
WaveMetrics; PO Box 2088; Lake Oswego, OR 97035; 503-620-3001
The following product description of the statistical software package SYSTAT comes from the on-line help facility on the UC Berkeley computer system "garnet.berkeley.edu". It was located in the help system under "help micros software mac systat".

**SYSTAT**

Product: SYSTAT  
Version/Release: Mac 5.1  
Vendor: SYSTAT, Inc.

Product description: SYSTAT is a comprehensive statistics package with a strong graphical component. Tools for statistical analysis range from regression models, analysis of variance, and discriminant analysis methods to time series models and factor analysis techniques. A wide variety of two- and three- dimensional graph types are provided. All tests and graphs can be invoked from pull-down menus and dialog boxes. SYSTAT is distributed in two versions, one of which requires a 68020/030 CPU and a math coprocessor. Compatible with System 7.0.

Hardware system: Macintosh Category IV (Plus or higher)  
Operating system: Macintosh OS 6.0.2 or higher

Additional hardware requirements:  
RAM: 2MB (4MB recommended)  
Disk storage: 4.3MB  
Math co-processor:  
Video display:

Additional software requirements: None

License or subscription fee: $75.00  
Fee type: Annual

Distribution procedures:  
Public file server (no distribution fee): No  
Fee for other methods of distribution: None  
Material to be provided by customer: Six DSDD 800KB 3.5" diskettes

Contact for ordering: Workstation Consulting & Demonstration Facility, drmicro@garnet or 642-8899.

Printed documentation: Included  
Online documentation: None

Comments: 1st yr cost $75, incl $50 for req'd manuals. Update $25/yr thereafter.

UCB August 21, 1991
2. Program Listing - Module Conversion Software

Following is an excerpted listing of the C software\textsuperscript{22} which I wrote to read the
data module files, create a log of defibrillator events, and reconstruct the
electrocardiogram data in one of several formats as an ASCII data file.

/*****/

main.c for Read Module Data

a program to read a .DAT module data file and produce an event log file

Edit History, Pending Modifications, Bug List

2/2/91 - created by porting PC code from FIB-TAB.C.
modified structure to emulate unix style command line interpreter.
added console printing of event log
added FIB-DATA.C code to select events for printing to EKG data file
modified ekg data file to include multiple strips

- 2/7 swapped bytes in header file
   added header line w/ version date 2/7/91

1.1 2/9 changed header line to version number
   changed byteswap() as separate function
   changed source code formatting to more compact style
   re-wrote most code using pointers
   added section and some line comments
   changed output format
   added add1 #defines for size specifications

1.2 2/10 reformatted comments
   finished rewriting w/ pointers, commenting
   added typedef boolean
   put increment word pointer w/ wrap into separate function
   added real-time clock printout and calculation
   got rid of unused variables
   added location of starting record to header

1.3 2/11 modified user interface from...
   3 files (in, log, ekg) on command line to...
   in from stdin, out from stdout, only prompt for ekg if name given

1.4c 2/25 reverted to 1.2 source code.
   added more tables of strings for decoding annotation record fields
   dropped some debugging data from output, using more translated strings
   sorted output records by time
   month adjusted by 1 to match date for call 44701190.dat

1.4d 2/27 fixed decoding of values in vote/analysis, analysis sum, and defibrillation recs
   changed time conversions to use 1/2 sec tics -- times still don't match

1.5 9/18 added "?" option to command line for summary of usage
   will print out help text whenever 1st argument is not an openable
   file name
   modified to allow naming of data file. log and ekg files still default to
   a.log/ekg as per documentation

1.5.1 9/20 changed mods comments to reflect new ideas on output and graph
   package interface (format of ekg file)

1.5.2 9/20 added label line as first line in ekg file with crummy code in
   ekg writing section

1.6 9/22 transpose output of columns/rows in ekg file.

1.7 9/22 added FFT routines, non-functionally

2.0 9/27 created module.c, module.h, messages.h
   added additional condition (scanf != 1) to exit program

\textsuperscript{22} © 1992 William B. Lober.
2.1 9/28 added current time printout in log file
9/29 added power spectrum calculations
reorganized comments/edit history
printed out event log and 3-D graphs of first seven EKG segs

2.2 10/2 began adding code to implement user-selectable output modes
options are selectable by run-time flags
write out different versions of output (user selectable):
   1. whole file vs. individual segments
   2. series of EKG units for each segment
      vs. triplets of (time w/in seg, EKG units, time of segment)
      (this allows relative display of segments in time)
automatically write and fft all EKG segments

2.2.1 10/3 changed order of columns in triplet output format
10/5 working on main.new - adding columns for FFT's

2.3 10/6 changed column flag to allow multiple output formats
      added single vector EKG output for IGOR, with
0's at unrecorded data points

2.4 10/29 changed command line input to accept a name, and use that
      name for .dat, .log, .ekg, and .fit files.
(2.5) 10/30 kludge for batch processing of files...
      uses hardwired array of strings

2.6 10/30 reads names from file called "names" for batch processing

2.7 11/11 converted EKG output to volts
      conversion factor is (0,255) = (0.0V, 5.0V) 1 unit = 19.600mV
      per phone conversation with Jim Angel 11/11

2.8 11/13 changed numrecs=1 from numrecs=i-1 in 1st assignment of numrecs
      this way includes next "illegal record" as a UNIT OFF marker
      commented out sort of records by time, as they should be
      written in a sorted order, and I am no longer excluding illegal records.
      added questions in CAPS to comment blocks regarding validity
      added comment to module.h re: record.mtime
      added (seconds from on-time) to log file
      versions 2.3-2.7 single column mode produced bad time scaling

2.9 11/14 suppressed EKG segments if "empty or being filled"
      stole time for "Illegal Record Type" from previous record,
      and changed string to "Unit Off" in messages.h
      WARNING - Only tested for Z_OUTPUT mode - tho should be no changes

(mod) verify data by comparison with Laerdahl Event Log
9/29 EKG data verified visually
verify times
(mod) add * to printout to indicate inclusion in Event Log
(mod) change input line to fit documentation in help string
(mod) find new graphing program
(mod) tidy up code in main.c, moving routines to other files.
(mod) turn into Mac Application

bugs

2.7 bugs
   Unit Off message not printed.
   Times wrong on 4470.dat final EKG's
   Times seem OK in Event Log?
works only with last episode found on a module. are there ever >1?
   seems like it - look at figures for buffer index of first record
wraparound works in theory... needs explicit test w/ wrapped data
   again, watch the buffer index and record count figs to find one
times seem wrong 0164 1051AM 12/28/90 vs. database 1051AM 11/28/90
```c
#include <stdio.h>
#include <stdlib.h>
#include <time.h>

#include "module.h"  /* data structures and module.c in defs */
#include "messages.h" /* message strings for the module routines */
#include "fft.h"      /* for fast fourier transform routines */

#include <console.h>   /* add command line emulation support */

#define V_PER_UNIT 0.019608 /* Conversion factor (see 2.7 comments) */

int command(char ***p);

/***
  Global Variables
***/

static char vers = " Version 2.9"; /* version string - printed in header */

static char name_buff[16];

ANN_REC records[350];

static int *prec_last, *prec_first;

static int ecg_tab[NUM_EKGS]; /* Table giving record number (in array 'records') for this ecg data segment */

static unsigned char data[MCM_ROM_SIZE+1];

main(int ac,char *av) {
    unsigned char *d_cur, *d_end; /* pointers for data buffer */
    int *aw_beg, *aw_cur, *aw_end; /* ptrs to allocation record buffer */
    boolean done, halvesecs_flg;
    int i, num_recs, tics;
    int rec_type, rec_index, ekg_index, ekg_old, seg_type;
    int index;
    int first_word, serial;
    float t;
    char t_buf[50];
    time_t cal_time, evt_time;
    struct tm real_time, *local_time;
    int index_count = 0, ekg_indices[NUM_EKGS], byte_count;
    float *fft_real, *fft_imaginary, *fft_output[NUM_EKGS], *fft_freqs;
    boolean prompt_flg = FALSE, column_flg = Z_OUTPUT;
    /* YZZZ_OUTPUT   define output formats where */
    /* XYZ_OUTPUT    x = segment time, y = time */
    /* Z_OUTPUT      and z = EKG/FFT value */
```
int ekg_times[NUM_EKGS];
float start_tic;
long int sample_time, strip_time;
char fname[13];

ac = command(&av);
if((names = fopen("names","r")) == NULL) {
    print_help();
    printf("Can't open 'names' as input file\n");
    exit(1);
}

while( fgets( name_buf, 16, names ) != NULL ) /* repeat program while not end */
    name_buf[strlen(name_buf) - 1] = '\0'; /* wipe out <NL> character */
    /* reinitialize stuff done above */
    index_count=0;

/****
Set-up Unix-style command line emulation in Think C
Open module data (in), event log (out), EKG data file (out), and fft data file (out)
from command line specifications. Exit if problems occur opening files.
****/
printf("Opening file %s\n",name_buf);

strcpy( fname, name_buf );
strcat( fname, ".dat");
if((in = fopen(fname,"rb")) == NULL) {
    print_help();
    printf("Can't open '%s\n',fname);
    exit(1);
}
    /* append ".log" to av[1] */
strcpy( fname, name_buf );
strcat( fname, ".log");
if((logfile = fopen(fname,"w")) == NULL) {
    printf("Can't open %s\n",fname);
    exit(1);
}
strcpy( fname, name_buf );
strcat( fname, ".ekg");
if((ekgfile = fopen(fname,"w")) == NULL) {
    printf("Can't open %s\n",fname);
    exit(1);
}
strcpy( fname, name_buf );
strcat( fname, ".fft");
if((fftfile = fopen(fname,"w")) == NULL) {
    printf("Can't open %s\n",fname);
    exit(1);
}
/*logfile = stdout;*/ /* temp redirection */

/****
read "n" bytes into data
****/
if((i = fread(data,1,MCM_ROM_SIZE,in)) != MCM_ROM_SIZE) {
printf("%error - number of characters read: %d != number asked for: %d\n",
i, MCM_ROM_SIZE);
exit(1);
}
fclose(in);

****
Starting after header, swap byte pairs til end of annotation records.
Required since data module stores bytes in reverse order.
****/
d_cur = data + HEADER_SIZE;
d_end = d_cur + ANN_TAB_SIZE; /* point to first word whose bytes don't
get swapped */
byteswap(d_cur, d_end);

****
Print the Header
****/
fprintf(logfile,"Event Log Generator - Mac %s\n", vers);
time(&cal_time);
fprintf(logfile,"File created at %s\n",ctime(&cal_time));
fprintf(logfile,"Annotation table for file: %s\n", name_buf);
fprintf(logfile,"Test words %x %x Nbr of shocks %d\n", data[0] & 0xff,
data[1] & 0xff, data[2] & 0xff);
fprintf("--Annotation table for file: %s\n", name_buf);
fprintf("--Test words %x %x Nbr of shocks %d\n", data[0] & 0xff,
data[1] & 0xff, data[2] & 0xff);

****
Find the end of the LAST EPISODE by looking for "Illegal Record Type"
(what if that was the first byte found? shouldn't we make sure we're in a section of valid records
first?)
First compute the boundaries (aw_beg and aw_end)
of the annotation record circular buffer
Then find a record of type "Illegal Record Type"
Error if no record of that type found.
aw_cur is set to first word, aw_end is set to point to last possible word.
****/
aw_beg = (int *)(data + HEADER_SIZE);
aw_end = aw_beg + ANN_TAB_SIZE / 2 - 2;
aw_cur = aw_beg;

while (aw_cur < aw_end) {
    if (*aw_cur & R_BIT_MASK) /* test for 1st word in a rec */
        if (((*aw_cur & R_TYPE_MASK) >> 11) == 15) {
            prec_last = aw_cur;
            break;
        }
    aw_cur++;
    /* next word */
}
if (aw_cur >= aw_end) {
    printf("%error - Can't find end-of-record block\n");
    exit(1);
}
}
3. Program Listing - Igor Procedure Software

Following is a listing of the procedure code\footnote{23} which I wrote for the Igor signal processing package. This code reads the EKG data and measures its amplitude and frequency characteristics, writing the results into an ASCII file.

```
************ EKG Graph Windows ************

Window WholeStrip(w) : Graph
    string w
    PauseUpdate; Silent 1   | building window...
    Display /W=(5,42,400,250) EKG as "Whole EKG Strip"
    Label left "Amplitude (Volts)"
    Label bottom "time (seconds)"
    SetAxis left 0,5
    TextBox /A=MC/X=18.0064/Y=44.586 "Complete EKG Strip"
EndMacro

Window ExpandedEKG(): Graph
    PauseUpdate; Silent 1   | building window...
    Display /W=(5,42,400,250) EKG as "Expanded EKG Strip"
    Label left "Amplitude (Volts)"
    Label bottom "time (seconds)"
    SetAxis left 0,5
    TextBox /A=MC/X=19.6141/Y=43.949 "Expanded EKG Strip"
EndMacro

************ Page Layouts for EKG data ************

| removed to save space...1/4/91

************ Statistical Functions for EKG data ************

Function PeakToPeak(w)
    wave w
    Variable biggest=0
    variable limit=numpnts(w)-3   | end when pointing to second-to-last element
    variable first=0, pt1=1,pt2=2   | the starting points for the search
    variable diff
    variable i=0
    do
        diff = YDiff(w,first,pt1)   | calculate the differences of the Y values at these points
        biggest = max( biggest, diff)   | always check
        if ( IsInflection(w,first,pt1,pt2) )   | is pt1 an inflection point?
            first = pt1
            pt1 = first+1
            pt2 = first+2
            if ( first > limit )   | the peak was at pt1
                break
            endif
            else
                if we can't increment, then we're done
                this is another end of the algorithm**********
    ```

\footnote{23} © 1992 William B. Lober.
\begin{verbatim}
pt1 = pt2
pt2 = pt2 + 1
if ( pt2 > limit+2 )
    break
endif
diff = Ydiff(w, first, pt1)
biggest = max( biggest, diff )
endif
while ( 1 )

    return biggest
End

Function IsInflection(w, p1, p2, p3)
    wave w
    variable p1, p2, p3
    if ( w[p1] > w[p2] )
        if ( w[p3] > w[p2] )
            return 1
        endif
    endif
    if( w[p1] < w[p2] )
        if( w[p3] < w[p2] )
            return 1
        endif
    endif
    return 0
End

Function YDiff(w, first, last)
    variable first, last
    wave w
    return abs( w[first] - w[last] )
End

************ Macros to Manually Process EKG data ************

Macro UseEKG(w)
    string w
    Prompt w "data wave:".popup WaveList("", ",", ",")
    PauseUpdate; silent 1
    wavenam = w
    if( exists("EKG") )
        Killwaves EKG
    endif
    Duplicate Sw EKG
    SetScale/P x 0.0.01,"" EKG;
    WholeStrip(w)
    ExpandedEKG()
    ShowInfo

EndMacro
\end{verbatim}
Macro ExtractSegment( num, x1, x2)
    variable num, x1,x2
    Prompt num,"Segment Number: ".popup "1;2;3;4"
    Prompt x1, "Starting Point"
    Prompt x2, "Ending Point (0 for 1024 element strip)"
    if x2 == 0
        x2 = x1 + 10.24  
        (* I use this to figure strip length directly, for consistancy *)
    endif
    string segstr=num2str(num)

    Tag EKG, x1, "Z07Seg "+segstr+" ("+num2str(x1)+","+num2str(x2)+")"
    Tag EKG, x2, "Z07Seg "+segstr+" ("+num2str(x1)+","+num2str(x2)+")"

    string pr_name="EKG_pr"+segstr
    Duplicate/R=(x1,x2) EKG $pr_name
    SubMean(pr_name,segstr,0 )
    string ms_name=pr_name+"_ms"
    string stats1="Peak To Peak: "+num2str(PeakToPeak($ms_name))
    Textbox /A=MC/X=23.5849/Y=17.8344 stats1
    Wavestats $ms_name
    Print stats1
    stats1 = "Range: "+num2str(V_max-V_min)
    Print stats1

    if( x2-x1 >= 5.12 )
        PSD(ms_name,3,2,segstr)
        SetAxis bottom 0,20
        Modify minor(bottom)=1
        Modify grid=1
        string tmp=ms_name+"_psd"
        Wavestats/Q $tmp
        string stats2="Peak Frequency: "+num2str(V_maxloc)
        Textbox /A=MC/X=23.5849/Y=17.8344 stats2
        Print stats2
    endif

    string off_name=pr_name+"_off"
    Duplicate $pr_name $off_name
    DeletePoints 0,1,$off_name
    string graphname = "rm_"+segstr
    Display $pr_name vs $off_name as graphname
    DoWindow/C $graphname
EndMacro

I------------------- Macros to Automatically Process EKG data(under development) -------------------

Macro LoadSingleWave( wname )
    string wname
    Prompt wname "Name of .ekg file:"

    string filepath="Point Beyond:Wave Analysis:EKG Files to load:"+wname+'.ekg"
    print filepath
    string wavename="w"+wname
    LoadWave/g=n=temp filepath
    Duplicate temp0 $wavenum
    Killwaves temp0
    Save $wavenum

66
Killwaves $w$wavenam
EndMacro

Macro LoadAWave($w$wavenam)
  string $w$wavenam
  Prompt $w$wavenam "Name of .ekg file:"

  string $w$filepath=diskname+":\Vfib Datasets:"+subfolder+";"+$w$wavenam+".ekg"
  print $w$filepath
  string $w$wavenam="w"+$w$wavenam
  LoadWave/$g$/n=$w$filepath
  Duplicate temp0 $w$wavenam
  Killwaves temp0
EndMacro

********** This is the main routine for processing EKGS. **********
The string for the sub-folder in which to look for the .EKG files
(which are used in LoadAWave) is passed as the global string "subfolder".
Then, cut and paste a routine which has repeated calls to ProcessWave,
to process the files in the subfolder.

Macro ProcessWave($n$, $x_1$, $x_2$, $w_1$, $w_2$)
  variable $n$
  variable $x_1$, $x_2$
  variable $w_1$, $w_2$

  l  wavenam, as a number...
  l  estimated start and stop of valid Vfib
  l  start and stop used in Wes's original measurements

  Silent 1
  errflg = 0

  l cumulative error flag

  l See if enough Vfib data is specified, then if enough of Wes's data is specified
  if( ($x_2$ - $x_1$) <= 3 )
    errflg = errflg + 1
  endif
  if( ($w_2$ - $w_1$) <= 3 )
    errflg = errflg + 2
  endif

  l Trim the end of the data strips by 1 sample.
  $x_2$ = $x_2$ - 0.01
  $w_2$ = $w_2$ - 0.01

  l *** Also, trim the $x_1$-$x_2$ interval if it is too big...

  l Choose the segment of interest
  variable $y_1$ = $x_1$
  variable $y_2$ = $x_2$

  l min of 1 sec req'd
  if( ($y_2$-$y_1$) >= 0.99 )
    l Convert wavenam to a string (without the initial "w")
    string $w$
    sprintf $w","%.4d","n$
    fprintf fptr,"%.5v","w$

    l Load the wave into memory and build its name
    LoadAWave($w$)
    string wavenam="w"+$w$

    l Process the wave (which writes out the results in the Hx file, according to errflg)
    AnalyzeAWave(wavenam,$y_1$,$y_2$)
Cleanup
   Killwaves $wavenam
   fprintf fptr,"v"
endif

EndMacro

Macro AnalyzeAWave(w, x1, x2) ...
   l combines UseEKG and Extract Segment code
   string w
   variable x1,x2
   Prompt w "data wave: ",popup WaveList("","","")
   Prompt x1, "Starting Point"
   Prompt x2, "Ending Point (0 for 700 element strip)"
  Pause Update; silent 1
   variable num=0
   l num = 0 is the indication that a temp file was created
   l by this routine (vs. ExtractSegment)
   if( x2 == 0 )
      x2 = x1 + 6.99
   endif
   string segstr=num2str(num)
   wavenam=w
   l shouldn't be req'd
   if( exists("EKG") )
      Killwaves EKG
   endif
   Duplicate Sw EKG
   SetScale/P x 0,0.01,"" EKG;
   WholeStrip(w)
   ExpandedEKG()
   ShowInfo
   Tag EKG, x1, "Z07Seg "+segstr+" ("+num2str(x1)+","+num2str(x2)+")"
   Tag EKG, x2, "Z07Seg "+segstr+" ("+num2str(x1)+","+num2str(x2)+")"

   string pr_name="EKG_pr"+segstr
   variable ptp = PeakToPeak($ms_name)
   string stats1="Peak To Peak: "+num2str(ptp)
   |Textbox /A=MC/X=23.5849/Y=17.8344 stats1
   Wavestats $ms_name
   Print stats1
   stats1 = "Range: "+num2str(V_max-V_min)
   Print stats1
   fprintf fptr,"%f%f%f%f", ptp, V_rms, (V_max-V_min)

   print numpnts($ms_name)

   V_maxloc = 0
   if( x2-x1 >= 2.56 )
      PSD(ms_name,3,2,segstr)
      SetAxis bottom 0,20
      Modify minor(bottom)=1
      Modify grid=1
      Wavestats/Q $tmp
string stats2="Peak Frequency: "+num2str(V_maxloc)
Textbox /A=M/C/X=23.5849/Y=17.8344 stats2
Print stats2
DoWindow /K psd_0
endif
if( V_maxloc >0 )
   fprintf fptr,"%f",V_maxloc
else
   fprintf fptr,"%s","*"  l is this %s really req'd?
endif

KillAllWindows()
KillWaves EKG, EKG_pr0, EKG_pr0_ms

EndMacro

********* Routines to batch process files in Datasets *********

Macro ThreeBatch()

**History of output runs from ThreeBatch**

1 0.9 dawn of time?
1 1.0 1/2 Whole Greg interval
1 1.1 1/4 Wes' interval for comparison of measurements
1 1.2 1/7 Greg interval for all calls, also, a zero for F_PEAK now always writes a "*"
1 1.3 4/11 Demo of Dataset 1A for greg
   string/G vers="1.4"
   string outfile="Three Batch Results "+vers
   string/G diskname="Oasis 105"
   variable/G errflg = 0  l cumulative error flag
   silent 1
   variable/G fpstr  l global pointer to text output file
   string filepath=diskname+':EKG Analysis:
   NewPath/O EKAnal.filepath
   Open /P=EKanal fpstr as outfile.
   fprintf fpstr,"Version %s on %s %s", vers, date(), time()
   fprintf fpstr,"%s\%s\%s\%s\%s","DAT_FILE","A_PTP","A_RMS","A_RANGE","F_PEAK"
   string/G subfolder="DataSet 1A (8)"
   Batch1A()
   string/G subfolder="DataSet 1B (22)"
   Batch1B()
   string/G subfolder="DataSet 2 (49)"
   Batch2()
   string/G subfolder="DataSet 3 (41)"
   Batch3()
   Close fpstr

EndMacro

Macro Batch1A()

ProcessWave(0798,1299,1313,1299,1306)
ProcessWave(1118,33,45,33,37)
ProcessWave(1709,47,56,47,49)
ProcessWave(3389,15,24,15,19)
ProcessWave(3567,34,41,0,0)
ProcessWave(3691,429,444,429,433)
ProcessWave(3693,6,17,6,10)
ProcessWave(5012,1505,1514,1505,1514)

EndMacro

Macro Batch1B()

ProcessWave(0846,3,18,3,10)
ProcessWave(1104,178,190,178,183)
ProcessWave(1183,2,15,0,12)
ProcessWave(1263,65,74,64,67)
ProcessWave(1381,61,72,60,65)
ProcessWave(1601,40,45,40,41)
ProcessWave(1948,80,92,80,83)
ProcessWave(2012,204,210,204,216)
ProcessWave(2106,24,42,24,33)
ProcessWave(2124,46,53,46,48)
ProcessWave(2202,44,62,44,52)
ProcessWave(2397,31,45,31,36)
ProcessWave(3536,18,23,18,23)
ProcessWave(3762,45,57,45,49)
ProcessWave(3870,9,0,9)
ProcessWave(4515,27,38,27,32)
ProcessWave(4673,87,98,87,92)
ProcessWave(5258,95,107,95,101)
ProcessWave(6011,20,29,20,22)
ProcessWave(6314,11,21,11,16)
ProcessWave(6668,59,71,59,65)
ProcessWave(6709,65,71,65,71)

EndMacro

Macro Batch3()

ProcessWave(0164,15,38,15,29)
ProcessWave(0246,58,71,58,65)
ProcessWave(0394,45,53,43,45)
ProcessWave(0429,15,27,15,21)
ProcessWave(0646,3,9,20,24)
ProcessWave(0687,97,105,0,0)
ProcessWave(0713,76,87,76,81)
ProcessWave(0807,66,75,66,70)
ProcessWave(0921,20,23,20,23)
ProcessWave(1221,41,53,41,46)
ProcessWave(1425,51,59,50,53)
ProcessWave(1485,72,81,72,77)
ProcessWave(1498,26,35,26,31)
ProcessWave(1547,22,28,5,9)
ProcessWave(2275,68,78,68,78)
ProcessWave(2422,54,68,54,59)
ProcessWave(2518,61,72,61,65)
ProcessWave(2638,95,97,95,97)
ProcessWave(2759,65,77,65,70)
ProcessWave(2866,315,327,315,319)
ProcessWave(2933,47,59,47,52)
ProcessWave(2990,64,74,63,66)
ProcessWave(3043,0,15,0,9)
ProcessWave(3152,39,51,39,43)
ProcessWave(3224,90,105,90,99)
ProcessWave(330,5,17,5,10)
ProcessWave(3539,68,80,68,73)
ProcessWave(3595,44,53,44,47)
ProcessWave(3710,68,77,68,72)
ProcessWave(3891,38,47,38,44)
ProcessWave(3976,83,93,83,88)
ProcessWave(4045,27,42,34,37)
ProcessWave(4237,51,65,51,57)
ProcessWave(4244,38,47,32,36)
ProcessWave(4460,112,117,107,112)
ProcessWave(4501,3,9,14,19)
ProcessWave(4577,11,21,11,15)
ProcessWave(4647,3,21,3,13)
ProcessWave(4767,14,26,14,19)
ProcessWave(4917,32,50,32,38)
ProcessWave(4927,21,29,21,26)

EndMacro

<table>
<thead>
<tr>
<th>Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>11. Change ProcessWave to take only one set of times and to take general EKG file names</td>
</tr>
<tr>
<td>12. Write module to read name and time from user</td>
</tr>
<tr>
<td>13. Write module to read names and times from batch file</td>
</tr>
</tbody>
</table>
4. Program Listing - SYSTAT Command Software

The following is the software\textsuperscript{24} I wrote for the SYSTAT command interpreter to recode the data and perform both graphical and statistical comparisons of the measurements both with each other and with outcome.

#####
Initial Recode Routine

takes dataset directly imported from FileMaker export
  and adds fields to scale Maxamp by 10 (to convert to mV.)
  and then take log base 10 transforms of all measurements
to get meet normality assumption for T-test, ANOVA

#####
DATA
USE "Oasis 105:Statistical Analysis:SYSTAT work:Vfib Systat"

let maxamp_s = maxamp/10
let maxamp_l = log(maxamp_s)/log(10)

let a_p_log = log(a_p)/log(10)
let a_r_log = log(a_r)/log(10)
let w_p_log = log(w_p)/log(10)
let w_r_log = log(w_r)/log(10)

let outcome2 = outcome
code outcome2/3=0 2=0 1=0 0=0.4=1

SAVE "Oasis 105:Statistical Analysis:SYSTAT work:Vfib Systat Logs"
RUN

#####
Graph all points, from Vfib Systat Logs
#####

EDIT "Oasis 105:Statistical Analysis:SYSTAT work:Vfib Systat Logs"

graph
select outcome > 0
ORIGIN -7.7 0.9
PLOT maxamp_s * w_p/ptp/SYMBOL= 2 FILL=outcome2 SIZE= 1 TITLE="Manual vs. DSP Measurements" HEIGHT= 3.50 IN WIDTH= 3.50 IN
ORIGIN 15.0 8.0
PICT "Oasis 105:Statistical Analysis:SYSTAT work:maxamp_s vs. w_p/ptp"

#####
Then draw circle excluding 7 outlier points, and
  save as Vfib Systat Logs Extract
#####

SAVE "Oasis 105:Statistical Analysis:SYSTAT work:Vfib Systat Logs Extract"
EXTRACT

#####
Now re-do previous graph as "Outliers", with previous layout
#####

\textsuperscript{24} © 1992 William B. Lober.
EDIT "Oasis 105:Statistical Analysis:SYSTAT work:Vfib Systat Logs Extract"
   graph
   select outcome > 0
   ORIGIN   -7.7   0.9
   PLOT maxamp_s * w_pip/SYMBOL=  2 FILL=outcome2 SIZE=  1 TITLE="Outliers" HEIGHT=  3.50
   IN WIDTH=  3.50 IN
   ORIGIN   15.0  8.0
   PICT "Oasis 105:Statistical Analysis:SYSTAT work:outliers"

    Pearson Correlation Coefficient of peak to peak amplitudes
    
EDIT "Oasis 105:Statistical Analysis:SYSTAT work:Vfib Systat Logs Extract Sorted"
   CORR
   PEARSON  W_PTP MAXAMP_S/PROB

    Prepare a By-Outcome2 dataset
    
SYSTAT
   SAVE "Oasis 105:Statistical Analysis:SYSTAT work:Vfib Systat Logs Extract Sorted"
   SORT    OUTCOME2

    Create density plots by outcome2 (Post process with McDraw)
    
EDIT "Oasis 105:Statistical Analysis:SYSTAT work:Vfib Systat Logs Extract Sorted"
   BY      OUTCOME2
   GRAPH
   DENSITY W_PTP/SMOOTH=NORMAL

    T-Tests
    
stats
   OUTPUT "Oasis 105:Statistical Analysis:SYSTAT work:T Test log"

   TTEST maxamp_s * DISCH$

   TTEST w_pip * DISCH$
   TTEST w_range*disch$
   TTEST w_rms * DISCH$
   TTEST W_P_PEPK * DISCH$
   TTEST maxamp_1 * DISCH$
   TTEST W_P_LOG * DISCH$
   TTEST W_R_LOG * DISCH$

    Scatterplots for bivariate analysis
    
graph
   ORIGIN   2 2
PLOTTING

```
PLOT w_range * w_prt/log=10 xlog=10 SYMBOL= 2 FILL=outcome2 SIZE= 1 TITLE="Log Range vs Log PTP" HEIGHT= 3.5 IN WIDTH= 3.5 IN ORIGIN 15.0 8.0 PICT "Oasis 105:Statistical Analysis:SYSTAT work:Log Range vs Log PTP"
```

```
PLOT w_rms * w_prt/log=10 xlog=10 SYMBOL= 2 FILL=outcome2 SIZE= 1 TITLE="Log RMS vs Log PTP" HEIGHT= 3.50 IN WIDTH= 3.50 IN ORIGIN 15.0 8.0 PICT "Oasis 105:Statistical Analysis:SYSTAT work:Log RMS vs Log PTP"
```

```
PLOT wf_peak * w_prt/xlog=10 SYMBOL= 2 FILL=outcome2 SIZE= 1 TITLE="Peak Freq vs Log PTP" HEIGHT= 3.50 IN WIDTH= 3.50 IN ORIGIN 15.0 8.0 PICT "Oasis 105:Statistical Analysis:SYSTAT work:Peak Freq vs Log PTP"
```

```
PLOT w_range * w_rms/log=10 xlog=10 SYMBOL= 2 FILL=outcome2 SIZE= 1 TITLE="Log Range vs Log RMS" HEIGHT= 3.50 IN WIDTH= 3.50 IN ORIGIN 15.0 8.0 PICT "Oasis 105:Statistical Analysis:SYSTAT work:Log Range vs Log RMS"
```

```
PLOT wf_peak * w_rms/xlog=10 SYMBOL= 2 FILL=outcome2 SIZE= 1 TITLE="Peak Freq vs Log RMS" HEIGHT= 3.50 IN WIDTH= 3.50 IN ORIGIN 15.0 8.0 PICT "Oasis 105:Statistical Analysis:SYSTAT work:Peak Freq vs Log RMS"
```

```
PLOT w_range * wf_peak/log=10 SYMBOL= 2 FILL=outcome2 SIZE= 1 TITLE="Log Range vs Peak Freq" HEIGHT= 3.50 IN WIDTH= 3.50 IN ORIGIN 15.0 8.0 PICT "Oasis 105:Statistical Analysis:SYSTAT work:Log Range vs Peak Freq"
```

---

Pearson Correlation Coefficients of amplitude measurements

---

EDIT "Oasis 105:Statistical Analysis:SYSTAT work:Vfib Systat Logs Extract Sorted"
CORR
PEARSON W_PTP w_range/PROB

CORR
PEARSON W_PTP w_rms/PROB

CORR
PEARSON W_PTP wf_peak/PROB

---

Scatterplots of Barton’s median frequency and my peak frequency

---

EDIT "Oasis 105:Statistical Analysis:SYSTAT work:Vfib Systat Logs Extract"

GRAPH
ORIGIN 2.2
PLOTTING

```
PLOT fm_wave * wf_peak/SYMBOL= 2 FILL=outcome2 SIZE= 1 TITLE="Median Frequency vs Peak Frequency" HEIGHT= 3.50 IN WIDTH= 3.50 IN
```

74
stats
CORR
PEARSON fm_wave wf_peak/PROB

####
Scatterplot of Median Frequency and peak to peak amplitude
####

ORIGIN 15.0 8.0
PICT "Point Beyond:Desktop Folder:Median Freq vs Peak Freq"

ORIGIN 2 2
PLOT fm_wave * w_ptp/xlog=10 SYMBOL= 2 FILL=outcome2 SIZE= 1 TITLE="Median Freq vs Log PTP" HEIGHT= 3.50 IN WIDTH= 3.50 IN
ORIGIN 15.0 8.0
PICT "Oasis 105:Statistical Analysis:SYSTAT work:Median Freq vs Log PTP"