An Expert System for the Treatment of Malaria

by

Catherine Maria Coury

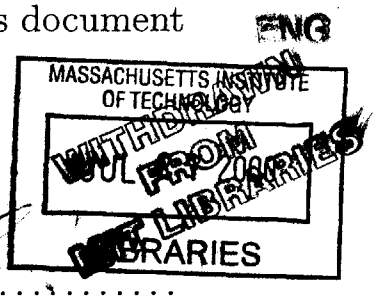
Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1999



Author .....
Department of Electrical Engineering and Computer Science
August 17, 1999

Certified by .....
Peter Szolovits
Professor
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

An Expert System for the Treatment of Malaria

by

Catherine Maria Coury

Submitted to the Department of Electrical Engineering and Computer Science
on August 17, 1999, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

This thesis details the creation of a prototype expert system for malaria treatment. Additionally, the prototype has the capability to interact with a novel mathematical model for the modeling of the interaction between the malaria parasite, the drug, and the immune system, as well as other modules; ultimately, this will allow it to give medication recommendations which take predicted parasite load into account. It will also have data collection capabilities, ensuring that the valuable epidemiological data collected by the system will be preserved.

Thesis Supervisor: Peter Szolovits
Title: Professor

Acknowledgments

Thank you to Cungen Cao, Jon Doyle, Hamish Fraser, Bill Long, and everyone else in MEDG; to Ong Lean Suan, author of the IMEX system; to Yemani, for his advice on malaria; to Ira Cooper, Chris Falling, Mark Krivan, and everyone else I forgot; and especially to Professor Peter Szolovits and Mojdeh Mohtashemi for helping me through this, providing knowledge and guidance, and everything else. Also thank you to Mojdeh for being the anonymous person behind the mathematical modeling.

Contents

Chapter 1

Introduction

Our aim is to develop a system to manage chemotherapeutic treatment of malaria that is also convenient to use. This thesis describes a prototype implementation. The system hopes to address two major problems. The first of these is the conventional method of malaria treatment and drug administration; malaria is endemic in many areas of the world, and while the drugs to treat it are inexpensive, they are typically either purchased over-the-counter or administered by untrained health workers. From this perspective, a system which prescribed medication doses would be quite useful; it would allow resources to be much better allocated. Furthermore, the system will be accessible all over the world via the World Wide Web.

The second problem addressed here is that of effective therapy. The program will make available a mathematical model; it is hoped that this model will improve therapeutic treatment of malaria. The model estimates the maximum pathogenic load of a patient, the time at which it occurs, and the optimal time to intervene. The resulting dynamics can then be compared to that resulting from the conventional scheme.

Chapter 2

Background

2.1 Databases and Data Collection

For most infectious diseases, there is very little data, either parasitological or immunological. The problem is not entirely that the data does not exist; rather, it is that the data is not being recorded in a fashion which leaves it available for further use. It is recorded on paper, or stored on computers which cannot access the Internet. Even if the computer storing the data has the ability to access the Internet, currently nobody receives this data. A system with a centralized database would therefore aid greatly in data collecting abilities.

Additionally, some of the data which is being searched for does not even exist at all. In impoverished areas, it is extremely difficult to take such data measurements as the parasite count in a person's blood. Therefore, in the few events where this data is taken, a way to preserve it is needed. Similarly, doctors who have the resources are more likely to take the data itself if they feel that doing so is likely to have a positive impact. If taken, the benefits of this data can then be brought to areas where neither the data nor the analysis is available.

2.2 Expert Systems

The term "expert system" is often used to refer to rule-based systems[10] which attempt to embody all of the knowledge present in a human expert of the same field. These systems specifically employ many techniques; two of these, known as forward chaining and backward chaining, are both dependent on a set of data. Forward chaining systems go "forward" from the rules, looking at all of the rules (often stated in a facts-derivable conclusion, or if-then, form), determining which facts are true given those that are known, continuing with these as facts, and continuing this process until they determine that the desired fact is true, false, or indeterminate (or after deriving all of the facts they possibly can). Backward-chaining systems, on the other hand, try all of the possible hypotheses, seeing if they can determine that the facts that support them are true. These facts are then taken as hypotheses, and the cycle continues.

These are not the only types of decision-making system. Tree-based systems can be used to tackle problems like games such as chess in which each possibility leads to several other possibilities - where the chart of possible moves is a tree. In addition, there are ways of pruning the tree - of determining that a particular branch does not contain the answer, therefore eliminating it and reducing your search time.

Another type of expert system has a probability based component; in addition to deriving a conclusion through such methods as described above, these assign a probability to that conclusion. For instance, the diagnosis module of IMEX[8] works like this; rather than saying that a patient definitely has malaria, it will announce that the subject has malaria with a CF (certainty factor) of x, where x is a value from 0 to 100 describing how convinced the system is of the diagnosis. The MYCIN system, described below, also has this characteristic.

2.2.1 MYCIN

The MYCIN[1] system was developed from 1972 to 1980. The goal was fairly broad: it was developed to provide advice on diagnosis and therapy of various infectious

diseases. MYCIN had a rule base and a backward-chaining engine; what was novel at the time was its probabilistic rule base. Also, it had an explanation facility much like that of Swartout's system, described later in this paper.

The actual results of the system were excellent; the work it did rivaled that of specialists. While it, as a prototype, was never clinically used, many systems based upon it have been.

2.2.2 Heart Disease Program

Dr. William Long, also of the MIT Clinical Decision Making Group, has constructed the Heart Disease Program[6], a system that assists in diagnosing more precisely the cause of a patient's heart failure. This system combines both a model and a probability-based rule-driven system; additionally, the program now has temporal-based reasoning; this allows it to take into account when different events occurred, drugs were administered, and other such effects; these combine to make it far more effective than a rule-based system alone.

2.2.3 Treatment Systems - Digitalis

Howie Silverman worked on a system for the administration of digitalis as part of Project MAC, the precursor to the Laboratory for Computer Science. Done in 1975, his system, ANNA[7], focused on the administration of digoxin, a preparation of digitalis. This drug, given to patients with chronic heart failure or arrhythmias, has a very low threshold range between an ineffective amount, a therapeutic amount, and a toxic amount. Moreover, this amount differs from patient to patient; a level that might kill one person is not enough to produce a therapeutic effect for another. Therefore, Silverman created this expert system to assist doctors in medicating patients properly. This system uses preexisting rule inference software in order to determine what questions to ask, and to determine if the previous medication level was insufficient, acceptable, or too high. To prescribe the actual doses, however, Silverman used pharmacokinetic calculations to augment the rule-based system; this took the system

beyond merely a program following a flow chart. This combination of a mathematical model and a rule-based system makes Silverman's system very similar in overall structure to that proposed; it also made his system quite powerful.

William Swartout's report, *A Digitalis Therapy Advisor with Explanations*[9], extends upon Silverman's system. This system, written with the benefit of a prototype of OWL (a language for knowledge representation), has the ability beyond Silverman's to actually explain its reasoning process - i.e., the user can type "Why?" at the prompt and receive an explanation of what the system is doing.

2.2.4 IMEX

To our knowledge, the work which comes the closest to the conventional portion of the proposed system is IMEX, the Integrated Malaria EXpert system[8]. (There may be a malaria expert system, possibly based on IMEX, operational in Cameroon; IMEX, however, is the most recent system which has been written up.) IMEX was designed by Ong Lean Suan, from the University of Singapore. She describes a system which is used for both the diagnosis and treatment of malaria in Malaysia; the system's recommendations are focused on malaria in that area. Her system is divided into four separate modules: the initialization module, which gathers preliminary information from the user; the diagnosis module, which concludes the certainty that the patient does or does not have malaria; the treatment module, which, given that the patient has malaria, prescribes an optimal treatment program; and the drug information module, which contains information about the various remedies available for malaria, such as treatment regimens and side effects.

The treatment module is the one most similar to this proposed application of our system. It takes into account the patient's form of malaria, statistics such as age and gender, factors such as pregnancy and any drug allergies, and how advanced the treatment and the disease have become. The program then goes through a purely rule-based process, proposing a drug, dosage and mode of administration. Also, it warns the user of possible complications to look out for.

IMEX does not have a model aspect; rather, it is a purely rule-based system. Like

the proposed system, IMEX was a prototype; the prototype was left with the drug center in Malaysia.

2.3 Rationale for Malaria as the Application

The initial application of this system, described here, is an expert system for the treatment of malaria during the transient stages of the disease. This thesis details that application, as well as many possible others. There are many reasons why malaria is a sufficiently important disease that it is worth this amount of study.

2.3.1 Endemicity

One of the major reasons why malaria is an important disease to study is its virulence; malaria is a disease of endemic proportions in tropical areas of the world. It claims 1.5-2.7 million lives each year[2], many of which are children and most of whom live in tropical Africa. In many of these areas, nearly all of the population has contracted the disease by the age of five. The most dangerous time period comes right when a person contracts malaria, before the immune system has had a chance to respond. The parasite rapidly multiplies in the bloodstream before settling down to a more stable level; it is in this period of immune suppression that the parasite multiplies to very high levels, and the risk of death or serious harm is the greatest. Therefore, this is the most crucial time to intervene; an optimal drug administration, even in the face of drug resistance, can get the patient through this time of crisis.

2.3.2 Current Drug Therapies

Many drugs are currently known to combat malaria. While these are amazingly inexpensive, and often available at clinics and even over the counter in embattled areas, effectively administering them can be problematic when differences between individual patients and factors such as side effects and drug resistances are taken into account. Therefore, the problem addressed here is that of prescribing an effective drug

regimen during the transient period of a patient's infection with the malaria parasite, when they are at the most risk. While many initiatives to control the malaria parasite are under way, that is all that is happening at this point - control. Eradication of the disease is currently impossible. The drugs reduce a patient's parasite load; this is one of the factors considered in the next chapter, as while it is not the only indicator, it is one way to determine the severity of the disease.

2.3.3 Parasite Count

The parasite count of a patient is a measure of the number of malaria parasites in their blood, per cubic milliliter. 100,000 to 200,000 is often a cause for alarm, and often accompanies a severe infection, although the parasite count itself is not the only factor. It is estimated that a count of 750,000 is perhaps the upper limit of what is survivable.[4]

Chapter 3

Demonstration

A system has been constructed that has a conventional expert system aspect, allows one to use a model for simulating the interaction of the malaria parasite with the patient, and has a data collection interface. These are all shown, in that order, in this section.

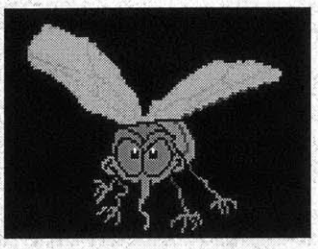
3.1 System Overview

The system starts with a welcoming web page. The first few pages allow the user to answer some basic questions; after that, he is then shown the results. Alternatively, he can access the data collection system from here; if he chooses to do this, he will answer questions pertaining to past results collected, and see the patient's parasite count and temperature as it changes over time.

In the conventional rule-based system, he is shown the results, both finding out what medication is recommended and seeing a graph of the patient's expected state. After this, he is given the option to continue into the experimental module. Here, he can get a dosage recommendation which, given the drug chosen, returns an optimal dose. (How an optimal dose is defined will be discussed later.)

The next section presents a more detailed system overview.

3.2 Screen Shots



Welcome to the malaria treatment system.

One purpose of this system is to select an antimalarial medication and prescribe a proper dosage, based on the guidelines for malarial treatment.
Another capability which the system has is on-line data collection. If you are here to provide data on patient parasite counts, please **click here**.

If you are interested in the first purpose, please provide some information about the patient.

Enter the patient's first name: Mark

Enter the patient's last name: Krivan

How old is the patient? 23 years, 1 months

Approximately how much does the patient weigh? 150

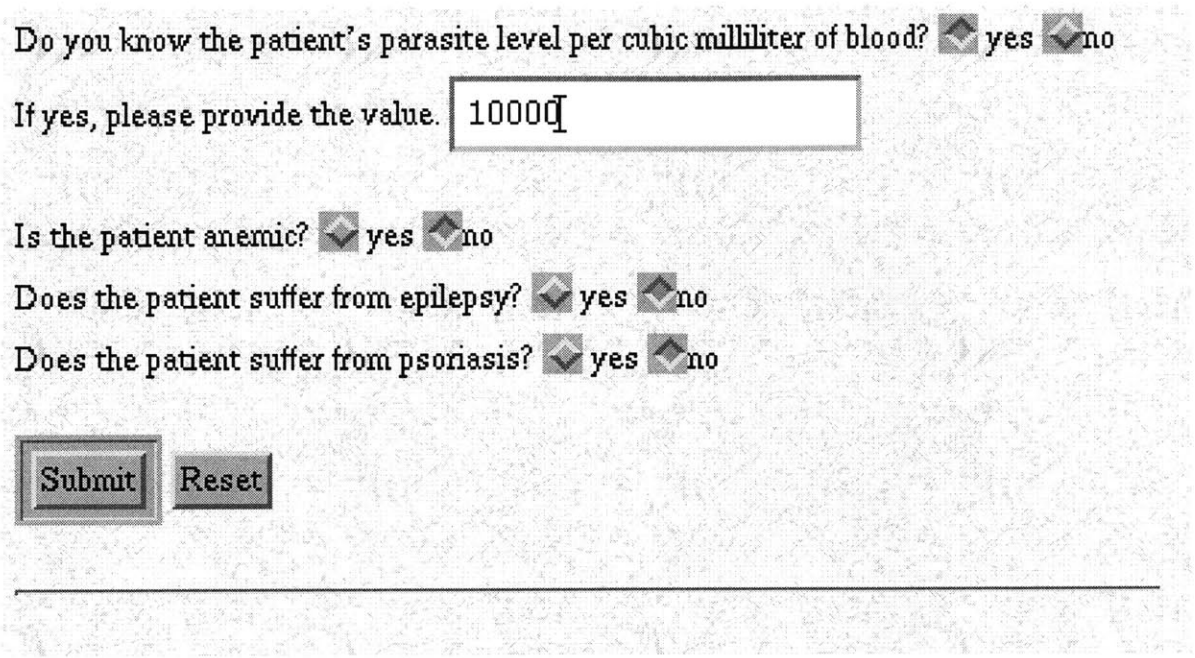
Is this in kilograms or pounds? kilograms pounds

What is the patient's temperature? 99

Is this in degrees Celsius or Fahrenheit? Celsius Fahrenheit

Is the patient male, or female?

Submit Reset



The first image is what a user initially sees upon starting the system. Here, and throughout this section, values for an example patient have been entered.

The values for parasite count and the corresponding temperature are real values.[4] The rest of the values are created so as to attempt to create a realistic person; in particular, the name is completely fictional.

The initial welcoming page explains the purpose of the system, and collects some fairly basic data which is needed of all visitors. Questions are designed so as to be simple as possible to read, understand, and answer. Also, note the data collection link in the first paragraph; this will be more thoroughly explored later.

Is the patient capable of taking drugs orally? yes no

Please check the drugs that are available to you.

quinine chloroquine Fansidar artemether

Please select the region of the world where you are located.

Other

If you selected "Other", please give that location. United States

---

When was this data collected? Jan 5 , 1999 12 : 01

AM PM

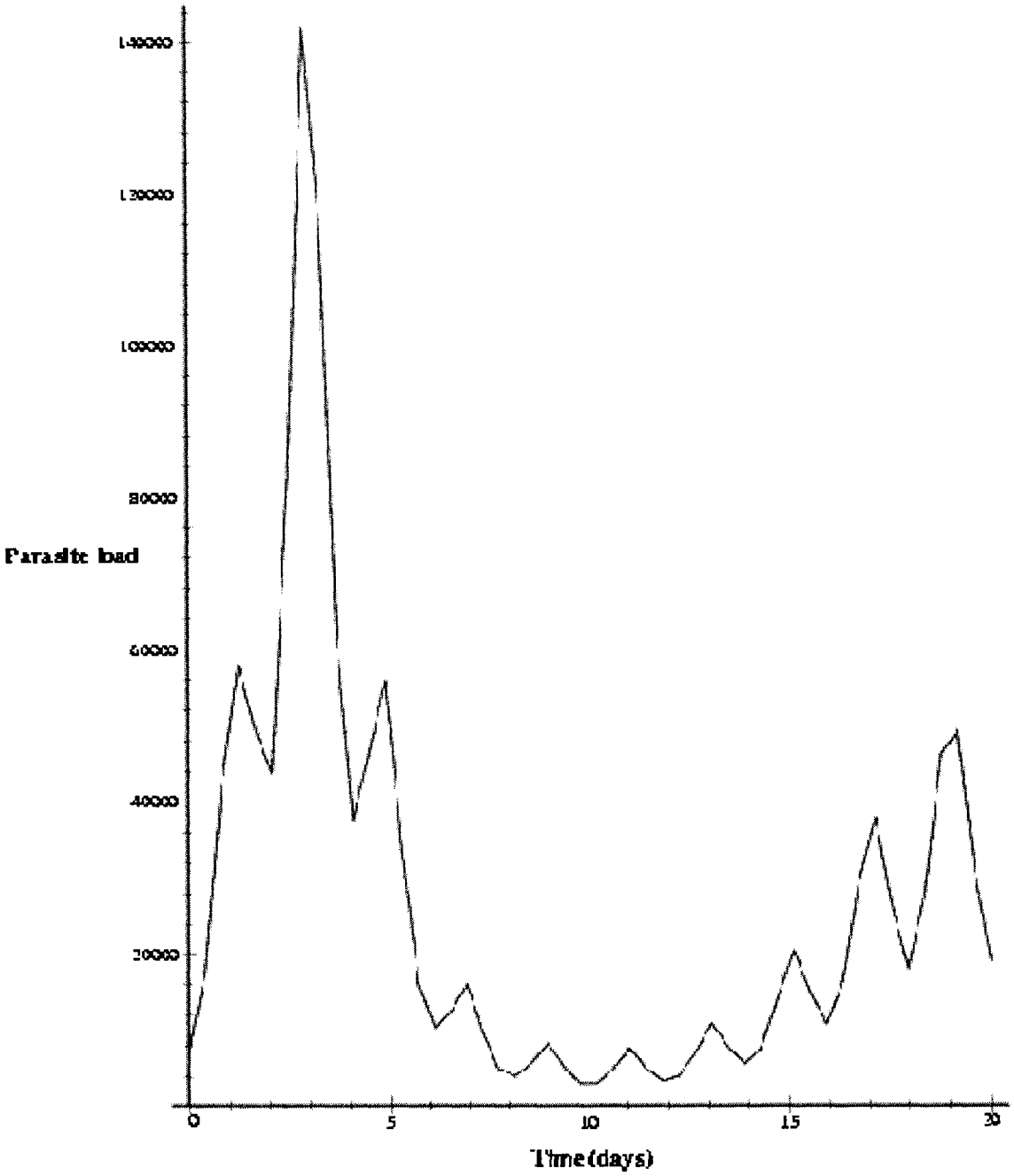
Submit Reset

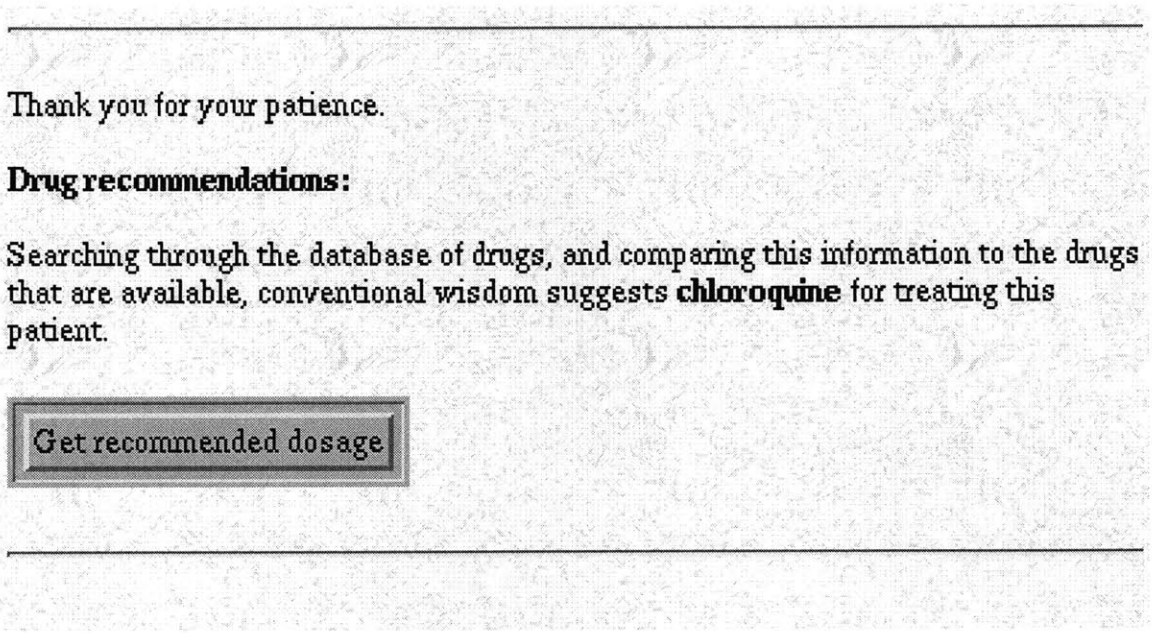
This is a graph of the patient's projected parasite level as time progresses, without any medication. The parasite level will peak in 3.08 days; the value in the person's blood at that time will be 143511 parasites per cubic milliliter.

After the conditions page, the system inquires about drug availability and location. Location is needed due to the various drug resistant strains of *P. falciparum*. The time is also asked; if more data is given in the data collection module, then this will be needed to track the patient's progress.

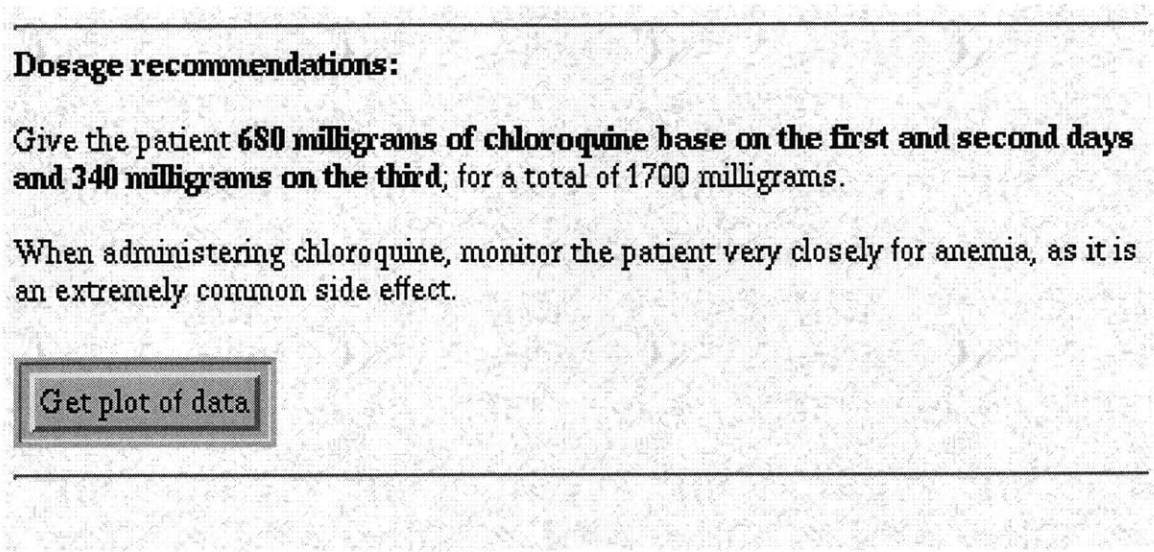
Now, the user is given the first of several pages of results. (The text above, along with the next plot, are all shown on one screen, along with some explanatory text; it is not shown this way due to size constraints of the thesis.) Here, you see the patient's predicted parasite load, and a graph of this over time. This plot is generated from the mathematical model, which is an integral part of the system. The maximum

pathogenic load and the time of its occurrence are estimated from the model. For example, for these dynamics, the maximum pathogenic load of approximately 143,511 parasites is estimated to occur at 3.08 days. Therefore, the optimal time to intervene is approximately 1 to 2.5 days after the initiation of infection. This is one of the novel features of this model, and will be further explained in section 4.4.1.





After this, the system then gives its drug recommendation. These pages has been made very short; buttons are provided on the bottom for the user to continue when he wishes.



Following the recommended drug is the recommended dosage of this drug. Here, it is highlighted in bold, so that the most important information can be easily seen by the doctor.

Below this, the program lists cautionary warnings about this particular drug. These will change, depending on circumstance; if, for instance, the doctor had previously answered "yes" to "Does the patient suffer from anemia?", then the likelihood of chloroquine's being chosen would have gone down. (See Appendix A for a copy of the actual algorithm.) If it had after this been chosen, a different and more severe cautionary warning would have been shown.

Other possible concerns:

Especially in hot climates, patients with malaria can often become dehydrated. This should be combatted, whether it be through simply drinking more fluids, or for children, oral rehydration solutions with extra glucose.

Fever can be a concern. This should be reduced, either by traditional methods such as attempting to keep the patient cool by removing clothes, sponging a large area of the skin, and fanning; or with antipyretic drugs such as acetaminophen.

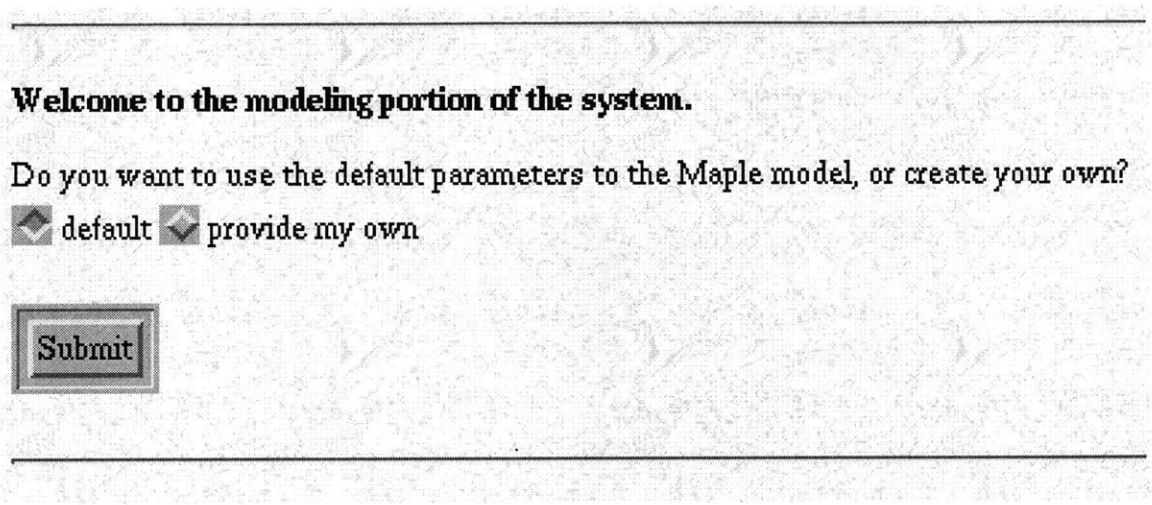
Another feature of this system is its experimental section, which is an experimental program that provides dosage recommendations based on a mathematical model done in the symbolic computation language Maple.

Would you like a comparative dosage recommendation? yes no

Submit

Here, the system conducts a wrapup of the conventional dosage portion. This is shown to all users, as it is relevant to all.

The system then begins its introduction to the experimental section of the program, with the Maple model. If the user replies no, they are given a thank you message; if they say yes, they are taken to the next screen shown here.



Here, the user is given an option as to whether to use the default parameters to the model (intended to take the chosen drug and calculate the best possible dose; possibly, ultimately, to take a list of available drugs and calculate the best drug and dose.)

The default parameters reflect an average person. These parameters take into

account such factors as the rate of decay and induction of the immune system, the strength of the initial immunity, reproductive rate of the pathogen, rate of removal of the pathogen by the immune system, time of immune system response, and initial pathogenic load. If for some reason the physician knows this information, they have the option to enter it; when these parameters have been more closely analyzed, this option will be more useful. For now, the example patient chooses the default option, which bypasses the data entry screen and goes directly to the results. Once again, despite the fact that they do not vary yet, the results are calculated each time.

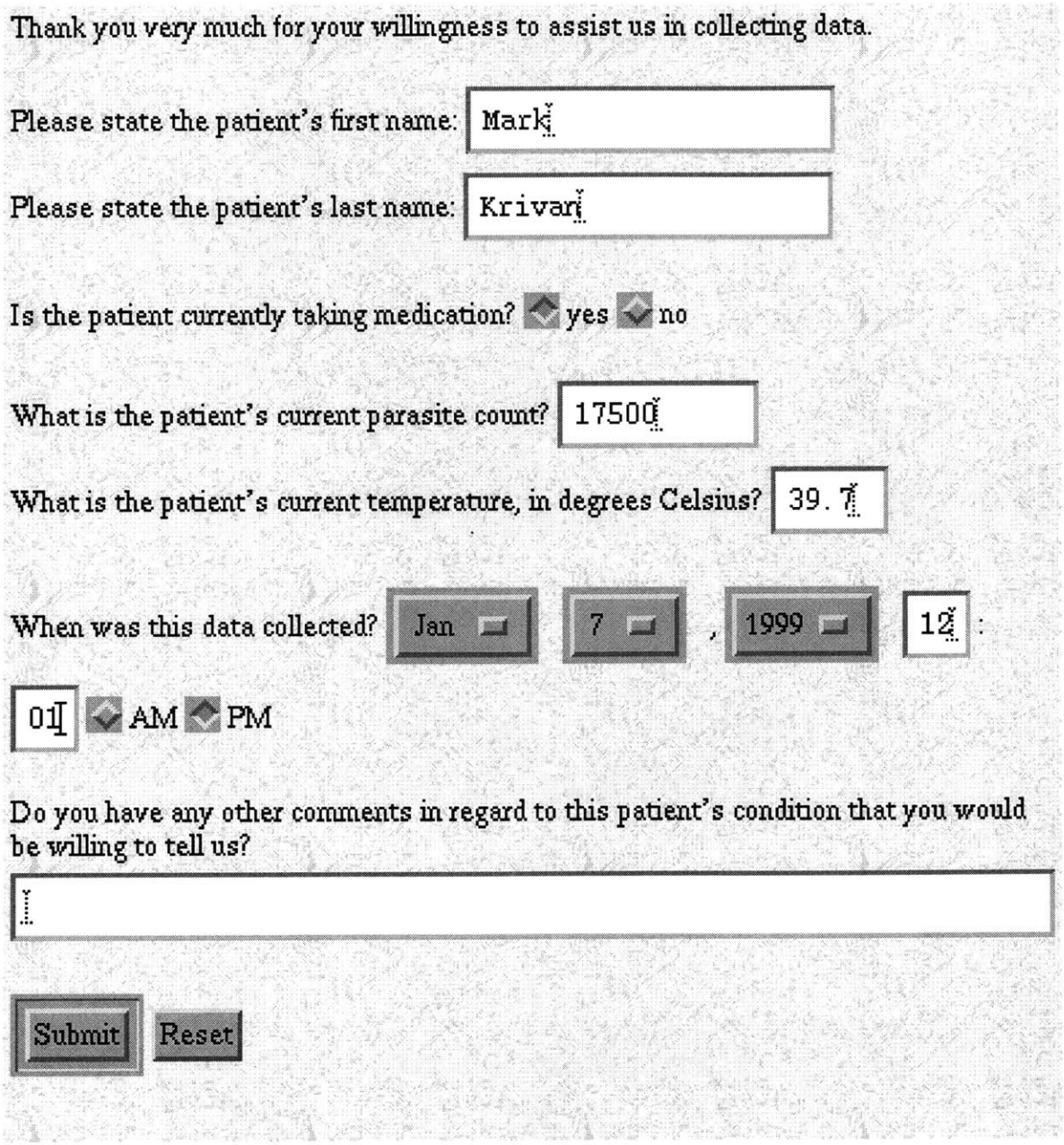
Thank you for your patience. The projected dose of **(drug name)** is 1 gazillion pounds a day.

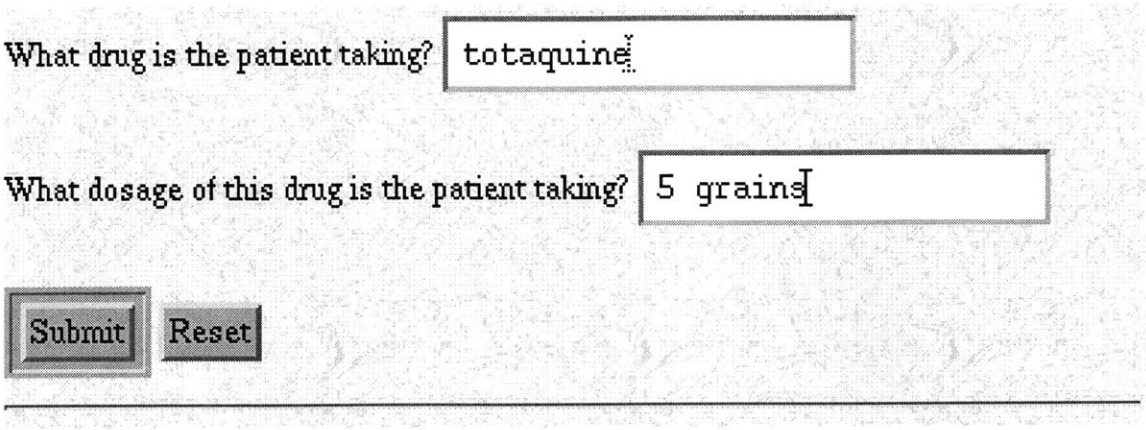
The graph below shows the patient's projected parasite level without any intervention (in blue) and with this dosage (in black).

Here, the doctor is told the results of the model - the recommended drug and dosage. (The example here, due to the fact that the model cannot make these pre-

dictions yet, is obviously fictional.)

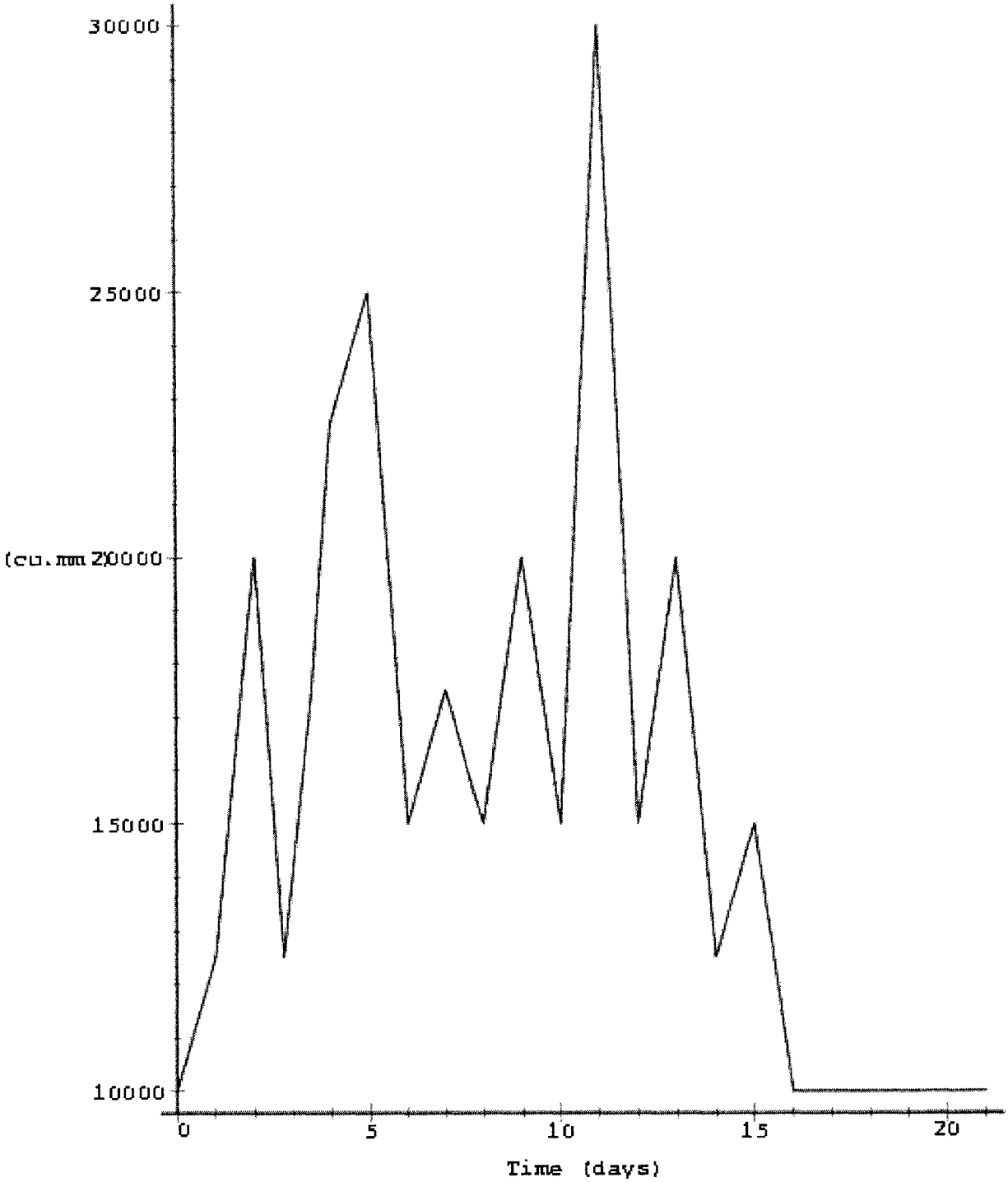
Then, the system displays the predicted parasite levels with the recommended medication, and tells the doctor about the data collection system. This is part of the same display as the above text.





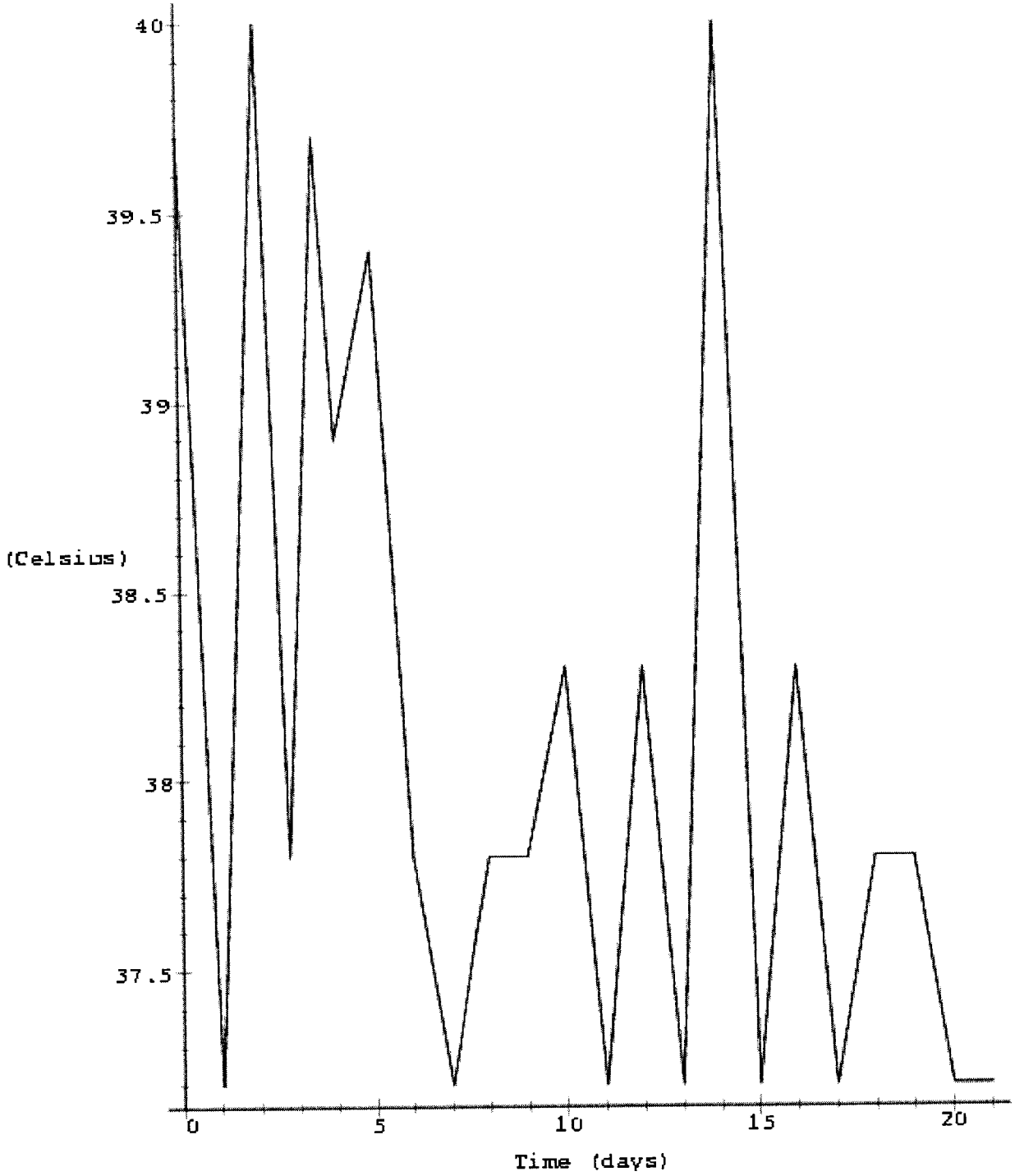
On the previous page, you will see an example of the initial screen in the data entry module. This screen corresponds to the pictures shown later; it is a point entered in the middle of this sequence. The times are relative to one another; therefore a start time of January 1, 1999 at 12:01AM was arbitrarily assigned.

If the doctor indicates that yes, the patient is under drug therapy, they are taken to this screen, where they can enter the drug and dosage amount which the person is taking. These values are stored for data analysis purposes, as are the values on the previous screen. In addition, the condition of the patient, as it has progressed, is plotted against time.


30000
25000
20000
15000
10000
0
5
10
15
20
Time (days)

Days since previous reading	Parasite count
0	10000
1	12500
1	20000
0.75	12500
0.75	17500
0.5	22500
1	25000
1	15000
1	17500
1	15000
1	20000
1	15000
1	30000
1	15000
1	20000
1	12500
1	15000
1	10000
1	10000
1	10000
1	10000
1	10000
1	10000

This chart, and the plot before it, along with some explanatory text, are shown on the same screen. Here, the physician is shown a plot and chart of all points that have been taken for this patient in the system. The first set before this text displays parasite count vs. time; the second, immediately following, displays temperature in degrees Celsius vs. time.


40
39.5
39
(Celsius)
38.5
38
37.5
0
5
10
15
20
Time (days)

Days since previous reading	Celsius temperature
0	39.7
1	37.2
1	40
0.75	37.8
0.75	39.7
0.5	38.9
1	39.4
1	37.8
1	37.2
1	37.8
1	37.8
1	38.3
1	37.2
1	38.3
1	37.2
1	40
1	37.2
1	38.3
1	37.2
1	37.8
1	37.8
1	37.2
1	37.2

Chapter 4

The System

There are two main components of the system which are currently available for use; the initial prescription model, which is constructed according to the guidelines[5] in *Bruce-Chwatt's Essential Malariology* and the data collection module. Additionally, the system has the capability to work with a Maple model; future work in this area is possible, and is described in the recommendations section.

4.1 Novel Characteristics of This System

First, this system differs from most other medical expert systems in that it provides treatment recommendations for malaria; this is only the second expert system to our knowledge (after IMEX) to do so. This is partially due to the areas where malaria occurs, as shown in Appendix E; many of them do not have the funds for such efforts. An example of this is seen in IMEX; despite the fact that Africa is a center of malarious activity, Suan's system was developed in southeast Asia and focused on the Malaysian area. Unlike hers, ours is applicable worldwide (that it is accessible on the web makes it available worldwide even now); however, it is not as specific about any one individual area.

Also, it differs in that it has additional features beyond those of a traditional expert system. It has a data collection facility; this is especially beneficial in regards to malaria, as currently very little data is available. Also, it can interact with a

mathematical model written in Maple; this function has been demonstrated with one particular model. It also has the capability to interact with other Maple models. Conceivably, this could be extended to allow it to work with any software package and any model, no matter what it is modeling. More details about possible methods for doing this can be found in the recommendations section.

The model for which it was initially intended, while not complete, will have the ability to predict a patient's parasite load over time, considering their immune system and any drugs which have been given. It can then recommend drug therapy that minimizes the maximum parasite load, perhaps giving a better recommendation in this respect than the traditional guidelines currently used by doctors.

While this model is not complete, it is nearly so. Additionally, the ability to hook into this model or any other exists.

4.2 Implementation Details

4.2.1 Prescription Module

The prescription module goes through the guidelines step by step and prescribes a drug. The actual algorithm which it uses is detailed in Appendix A, and an example record, as well as the format in which the data is stored, is in Appendix D. Again, these guidelines were derived from *Essential Malariology.*

The four drugs chosen were selected on the advice of a malariologist as the most likely to be available and representative of the different drug categories.

The algorithm was written using a point-based system; each drug is assigned a certain number of points based on various factors, and the drug with the most points is the drug that is chosen. This was selected as an extremely simple way to make what could be up to a four-way comparison. Additionally, it is a good way to reflect exactly how important each factor is; if there is one reason to choose a drug, and another against choosing it, for instance, this allows easy resolution of these conflicts.

4.2.2 File Formats

The model needs to have access to the parameters for it, as determined by the rest of the system, at the time Maple is run. Rather than placing the values in a file with perl, the language in which this program was written which will be discussed in more depth later, and then having Maple open and read that file, the values are simply inserted into the Maple file itself. This saves a lot of work for Maple, and is very little additional work for the perl module, as text processing abilities are a strength of perl.

4.2.3 Repeat Patients

When the system is told that a patient has been here before, it goes through the permanent database and finds all patients by that name. If no patient by that name is in the system, it is then forced to treat this person as a new patient. If, however, this patient does exist in the database, the program then returns to the user and confirms this with them. It does this by presenting a list of everyone with the same name as the patient in question in the form "Cathy #1 has been found. She is 23 years old and weighs 60 kilograms." From here, the clinician is given the opportunity to confirm which, if any, is the person to whom the clinician is referring. This section has been commented out of the final draft; however, it has been tested.

4.3 Conventional Drug Treatment

The system has a rule-based component. This is used when initially asking the user questions; the program needs to be able to know which questions to ask (i.e. have rules such as "if the patient is male, do not bother to ask if they are pregnant or lactating"). Once the information has been gathered from the user, it will be fed to both the mathematical model described above and a point-based prescription system.

4.4 Maple Interface

Another aspect of this program is its interface with the symbolic computation package Maple. This provides a facility such that you can connect any Maple model and work with it in the program. The possibilities of non-Maple models are discussed in the recommendations section. There is a section that allows you to say the location of the model on the server, what parameters you want to input, and then respond to its queries over the Web and have it return to you the values. This also makes Maple accessible over the Web. However, as the program has to be on the server, this is not anticipated to be a violation of Waterloo Maple's copyright.

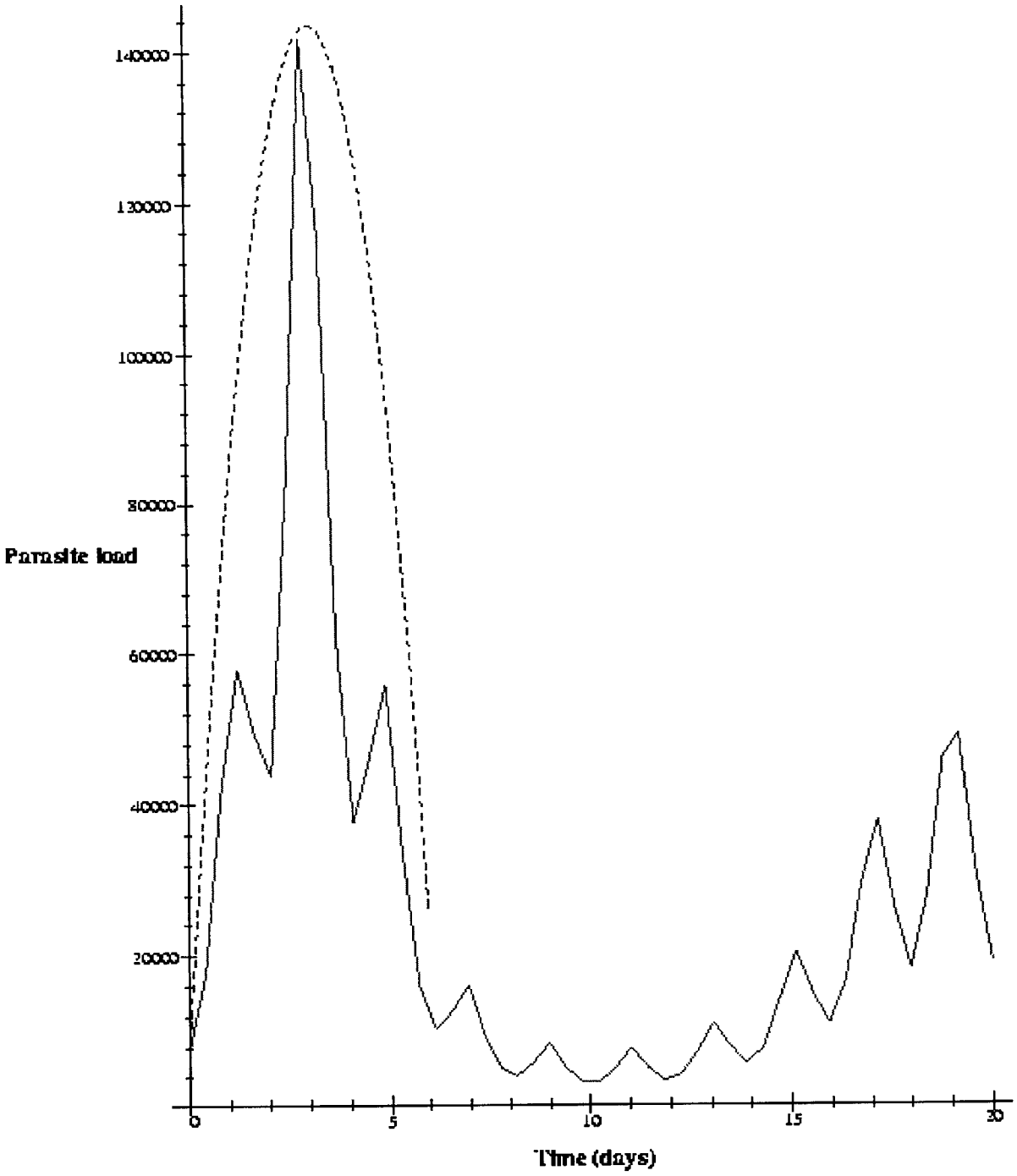
In order to add another model:

- For one run: Simply uncomment the code that provides this facility (in the test.html and options.pl files) and possibly make a few alterations in order to update the code base.
- For multiple runs: Create a Maple file (described above) and change all references to inittemplate to the new file. Alternatively, you can change the name of the file to inittemplate, the name previously used, and move the old file elsewhere.

4.4.1 Proposed Math Model

The proposed system will be working with an original mathematical model. This model will take in some patient-specific immunological and parasitological parameters. From there, it will prescribe an optimal chemotherapeutic regimen for the patient.

The models, both that used in the traditional portion, and in the experimental portion of the system, are innovative in and of themselves. Given the immune system-pathogen dynamic, these can estimate the height and time of the peak of the parasite load, and the time to intervene.


140000
120000
100000
80000
Parasite load
60000
40000
20000
0
5
10
15
30
Time (days)

As is seen in the plot on the previous page (taken from the example patient of Chapter 3), there is a dotted line. This dotted line is a parabola, projected onto the graph. By finding the peak of this parabola, the height of the crisis can be determined. Finding the time value of this gives the time at which the height of crisis will occur. Examining the parabola confirms the program's estimation of a peak level of 143,511 parasites after 3.08 days. The optimal time to intervene in this infection can also be seen; the approximate range of 1 to 2.5 days is determined by taking into account the two treatment factors described below.

The values plotted are found by modeling the interaction between the malaria parasite, the patient's immune system, and the drug which is given to the patient. Therefore, it can recommend an optimal drug regimen based on potency and half-life of available drugs.

There are two factors which become important when determining the optimal method of treating a patient. The first of this is minimizing the pathogen load; pathogen load per cubic milliliter of the patient's blood is a determining factor of the severity of an infection. However, the pathogen level cannot drop too low; if it does, the patient's immune system is tricked into thinking that the parasite is gone. It then stops fighting against it, and the parasite is allowed to multiply again. Therefore, the model aims for a "safe" threshold on the parasite load. Preliminary results from the model suggest that the dosage regimens it prescribes, which are lower than traditional amounts of the same medications, push the height of the crisis to occur chronologically later and result in a lower parasite count than that predicted for conventional prescriptions.

4.5 Data Collection

A goal of this system is to provide a simple, easy-to-use, integrated interface for data collection. It stores the data from the initial consultation; additionally, if the clinician is willing, he can input data on the patient's progress, such as further parasite counts and temperature readings. The system will then take this data and make it accessible

for medical research purposes. The data is stored on the server; therefore, interested parties can collect, or request, the data. It is also possible to email the data to researchers at any location, and code to do this was tested; many other options are available, as well. Currently, however, it is simply stored on the server.

The system stores records of nearly all of the information it is given, and all conclusions that it derives on a permanent database on the server. Currently, it does not store model parameters, as the model with which it will work varies and so will the number and meaning of the parameters. Therefore, it was deemed not meaningful to store this data. The rest is stored in order to preserve as much data as possible for further research; in a system dedicated to data collection, preservation of all data that can possibly be useful is an ideal feature.

4.5.1 Data Acquisition

In addition to that data collected in the data module, the system has already collected information from the patient's initial visit to the system. This is incorporated into the charts shown and data presented in the data collection module. This is only done the first time the data collection module is accessed in regards to a patient. As a design decision, it was deemed that checking only the first time was the most reasonable thing to do. This assumes that they do not go through the initial recommendation module after they go through the data collection module; while they can, that data point for that patient will not be counted.

Currently, if a doctor enters multiple readings for the same patient, only that which is entered the latest is kept. This is also a deliberate design decision. Keeping all of the points is suboptimal, as having two or more different values for data such as temperature and parasite count, which claim to have occurred at the exact same time, is nonsensical. Therefore, the system is presented with the dilemma of which one to keep. Keeping the last one allows for errors in data points to be corrected by the doctor; they can simply enter a new point, which is simpler than an error collecting facility. Additionally, if the physician inadvertently enters a data point twice, the last point is a more logical one to keep.

Nothing in the database can be changed. This is deliberate; as the points in the database are supposed to be a recording of fact. (Nothing in the database is actually recorded until the person reaches the alive.pl file, where a drug recommendation is given, which is well past the information-entering phase. Additionally, due to the nature of the Web-based program, the person can change their answers on a page at anytime while on that page, or if they need to change prior answers, by hitting their browser's Back button.) This also is an extremely simple way to prevent people from changing things which they should not be able to, assuming the user is not malicious; the problems of a malicious user are discussed in the web security section.

4.6 System Requirements

Currently, the server has a Pentium Pro 200 processor with 32 megs of RAM. This installation, on the computer on which the system was developed, can be found at the URL http://krivan.lcs.mit.edu/test.html. The performance is certainly acceptable here. The system is not bound by processor speed, although it will certainly run slower on a slower platform. However, there are a few requirements for any machine which wishes to act as a server.

1. The system needs a web server. krivan.lcs.mit.edu uses Apache, which is freely available for many platforms.

2. Maple on the system; alternatively, a copy of whatever third-party package the system interacts with, if any, will be needed.

3. perl. It is freely available via download.

4. An image converter, to make PostScript images, the only form that Maple can output, into images, such as .gif and .jpg images, visible on the Web without alterations to the user's browser. This implementation uses ImageMagick, a freeware program for NT and other form of Windows to convert the images to .gif format; ghostview, another freely downloadable utility, is needed for

this function within ImageMagick, and is an important factor in many such programs.

5. Depending on the environment, a few minor modifications to the code base may be needed. The system needs to know where Maple and the image converter are. Also, some of the copy/move/delete commands are those in DOS, and would need to be altered to run under Unix, in order to achieve the desired behavior.

To be a client, all you need is an Internet connection (or connection to the server some other way; for instance, any network supporting TCP/IP) and a web browser. It can be installed on a machine using that machine as both the client and the server; this allows it to be used on a computer with no Internet connection whatsoever. However, then some of the data collection capabilities are lost until and if the computer is connected to the Internet again; while it can still collect the data, the data cannot be sent anywhere.

Also, some form of Windows is recommended for the server. The system was actually originally designed for Unix. However, difficulty was encountered getting the Unix version of Maple, when used as a command line program, to write PostScript files correctly. The design of the initial model required this to happen; due to the fact that you want Maple has to shut itself down completely, not just close the worksheet, and there is no command to let it do so in the X or Microsoft Windows versions, the command line version functionality is a necessity. Therefore, Windows was used. As an additional side benefit, Windows is perhaps the most widely used operating system, especially in the non-collegiate environment. Therefore if someone doesn't want to use the main installation on krivan.lcs.mit.edu, they are more likely to be able to run the program.

Unfortunately, tests have not been possible with many users attempting to access the system at one time. (The system passed tests with up to three.) Naturally, performance will be affected; as more people are in the system, specifically the portions requiring Maple, then the longer it will take. Due to Maple, and the interface between Maple and the expert system, it is impossible to leave the connection open; different

connections for each user must be made. This is not anticipated to be a problem in Windows Maple; the worst-case scenario, that where Maple can only work with one user at a time, will result in the users' requests being serviced in the order received.

An example of the data storage, as well as the format, can be found in Appendix D.

4.6.1 CGI

All of the pages in the system, except the introductory ones which are written in straight HTML, are implemented as CGI scripts; CGI is perhaps the convention most suited to interactive Web applications. There are very few other feasible alternatives. HTML itself is not sufficient to perform all requisite features. Java is a less suited language to the task at hand. Additionally, it requires servlets (which are much more complicated then necessary) to perform the required tasks, unless you wish to use the CGI convention with Java, which is very uncommon and as such not a well-developed area. While all of these points could be explored, it was deemed unnecessary. As each script acts independently, state is not saved from one file to the next. This was counteracted by using hidden fields; while the user does not see the data, the CGI script passes it from one screen to the next.

4.6.2 Security

The entire script runs out of cgi-bin, which brings up cgi security issues. There is a certain amount of insecurity any time that you place an application on the Web. However, there is a certain amount of safety we can incorporate. An example of this is processing strings so that a user cannot run commands which you do not intend on your system; for instance, if a patient's file, stored under their name, were to be deleted, a patient named "..*" would cause the files in the directory above that which stores the file to be deleted. This has been implemented.

Access to another patient's record, to a certain extent, is possible. A malicious user could get it to graph someone else's temperature and parasite data against time, by

giving their name and a dummy point. In a real-world system, this is not necessarily acceptable. In a prototype system, however, it is more reasonable than asking a doctor to remember a pass phrase or some other bit of information that only they would know about each patient. There is no way to get at the rest of the record that is known at present.

4.6.3 Perl

"Perl is a general-purpose programming language invented in 1987 by Larry Wall. With over one million users worldwide, it has become the language of choice for World Wide Web development, text processing, Internet services, mail filtering, graphical programming, systems administration, and every other task requiring portable and easily-developed solutions."[3]

It was chosen because of its excellence for text processing, which this program heavily depends on. Also, for CGI scripting, it is one of the preferred methods; there are many archives of free CGI scripts on the World Wide Web to provide examples and inspiration.

In the event that the program is run on multiple servers, the databases will have to be combined in some fashion. The database is a permanent one; the program appends to it and reads from it, but never erases it. (Currently, any changes need to be made by a user with write access on the server. In the future, a method of altering the database might be useful; the actual utility of this, however, will be further discussed later.) In this database, the patient's name and some vital statistics are stored. Additionally, the previous recommendations are stored here.

4.7 Web Design

4.7.1 Rationale

The World Wide Web was chosen as a medium in large part to promote accessibility. A user with Internet access does not need to have Maple on his computer; he can run

this program, which has access to Maple for the purpose of the calculations needed. However, the prototype is not at this point fit for medical usage; as a prototype, of which all the pieces are not complete, it is not sufficiently thoroughly tested. Finally, the Web has a clean, simple interface. A lot of people understand it and are familiar with it, and writing applications for it is a quite simple, well-charted area, allowing energy to be devoted elsewhere.

4.7.2 Data Entry

Currently, if a second patient is entered with the same name as a preexisting patient, their data are intermixed. The chances of this occurring, while not high, certainly exist, especially if the system begins to see widespread use. Therefore, while this is acceptable behavior in the prototype, possible corrections for the future are discussed in the recommendations section.

4.8 Methods of Evaluation

4.8.1 The Whole System

The system has been tested, both by myself and with more and less experienced users. Thus far, those sections available appear to work; however, this cannot be guaranteed. (Additionally, to a certain extent the system is not complete, and as such, will have to be tested when it is completed.)

4.8.2 Malaria Treatment Prescriptions

Due to the time, space, and economic constraints involved, this is a prototype system, rather than a full-fledged tested system running on laptops in Africa. As such, total evaluation is impossible; this situation cannot be completely simulated with the resources currently available.

However, there is some testing that can be conducted. The recommendations of the system can be shown to a malariologist, or other skilled person, who can determine

if they are reasonable. Similarly, one can find a database of doctor recommendations and run the system on those cases, and see how close the system's recommendations come to those of the doctor. Some partial recommendations, for instance, are inherently available.[4]

Not in this prototype (in addition to the impracticality of this, the prototype is not sufficiently close to finalized to make this worthwhile), but someday, the system can be actually field-tested and taken on laptops in remote locations. First, let doctors use it; see how it compares to what they actually decide. Then, after it passes this phase of development, perhaps it can be put in the hands of those less knowledgeable about malaria such as field workers. This is where the system can do the most good - in the hands of those who do not have the knowledge that it embodies.

4.8.3 Dictates of the Model

The theories of drug resistance and such are even harder to test. The possible treatment programs which it generates can certainly be at least sanity-checked. This can be done by having a more skilled expert in the field look at the dosages prescribed; alternatively, it can be done by comparing the dosages to those given out by the more conventional expert system aspect of the program. (Note however that these dosages are not supposed to be conventional; as such, this may not be the fairest comparison.)

Chapter 5

Recommendations

5.1 Future Algorithms

In the future we can envision a more complex algorithm that takes previous data (which we do have) into account - that uses this information for more points on a proverbial curve, better enabling the model to predict where on the curve the patient will be in many hours. While doctors can subconsciously do this, no such algorithm has been codified to our knowledge.

In addition, one of the possibilities for the more conventional model is pharmacokinetic calculations such as those in the digitalis model. Nothing like this has yet been implemented in this system.

Currently, the system has a facility for developers in which they can dictate parameters to a Maple model; the system will walk them through this and have that model run, rather than the conventional drug-resistance model. As this is commented out, it will need some updating, due to further system evolution, before being put into use again. Without this parameter translation, the model cannot be adapted to different people correctly. As such, it cannot reflect individuals.

5.2 Updating the Innovative Malaria Model

For a disease as widely spread as malaria, extremely little is known about its dynamics. It is an incredibly complex disease which is not well understood. The goal of the data collection is to help gather information about the dynamics of the disease. (This data can also help determine the correct parameters for the model; as such, it is needed for updating the model.)

The model does need some work, and that would be a recommendation. The parameters need to be better understood. Additionally, the interface of this model with the rest of the program is not perfected. The model's method of acquiring data is not finished; while there is a location for this in the main program, it is not perfected. The explanations for the parameters do not make perfect sense; only when they are better understood can more precise explanations be written.

5.3 The Interactive Model

Part of the flexibility of this research comes into play here; an entirely different idea is being explored - that of allowing a clinician to come up with different dosage ideas, perhaps guiding them to a few possibilities so as not to make the search space as daunting, and allowing them access to the model of the malaria parasite dynamic in the person's body. This access will allow them to see what is predicted with the recommended medication dose, with no dose, and with other doses which they might happen to think up. In order for this modeling to be patient-specific, the person-specific parameter problem described above needs to be solved. However, doing it with the generic person still has great merit.

The initial capability for this is already in the system - a first cut at this (the first iteration, without projecting the different dosages) is what is shown to the user before the recommendations are given to them. Additionally, this reads the height of crisis, defined as the peak of the parasite concentration in the person's bloodstream, and on which day and when this will occur, and reports these to the user. (Despite the fact

that the plot is currently the same every time, it does generate the plot individually; in a more advanced version, these plots will differ.) Iterating through this process, and continuing to report back to the user with the new information, is the idea which is described here.

5.4 Data Collection Improvements

The problem with data confusion between two people of the same name stems directly from the fact that the data is indexed under a person's name only. As such, it can easily be indexed with another piece of information as well; in fact, when the module to identify if people had visited the system before was active, it used the person's age and weight. The prototype, which does not see heavy usage, is perfectly fine in this stage; however, additional indexing information is needed if the system is to see serious use.

Some additional security tests need to be conducted. Like any other Web application, this one has a lot of possible security holes. While any security holes that a malicious user could exploit are not immediately known, additional testing by those more skilled in this area would be wise.

5.5 Platform Independence

The attempt to remove platform independence was not completely successful; this will be difficult without a "setup module", where the user provides the correct command paths, due to the dependence on Maple. Alternatively, different versions could be provided for standard Maple installations in DOS, Windows, Unix, and so forth. This would also fix the dependence on system utilities; either specific image conversion programs could be built, or a freeware program appropriate to each platform could be included.

Chapter 6

Conclusion

Here, we have presented the design of an expert system for the treatment of malaria. This is a prototype system; while it is not perfect, it illustrates many important principles.

- It deals with the problem of malaria; this in and of itself is rare.
- It provides an interface for data collection, something which is lacking in malaria studies.
- It can interact with a mathematical model, and provide additional information in this fashion, in addition to serving the functions of a standard rule-based system.

Appendix A

Drug Selection Algorithm

A.1 Drug Selection

The drug selection algorithm is a points-based algorithm, derived from Bruce-Chwatt.[5] The algorithm is detailed here.

Starting point values: quinine = -99999; chloroquine = -99989; Fansidar = -99997; artemether = -99994.

This indicates that, all other factors being equal, the order of choice is chloroquine - artemether - Fansidar - quinine. In the cases where chloroquine is effective, it is in fact the best drug to handle malaria.

For each of the drugs that the clinician has indicated they have on hand, add 99999 to its total number of points, thus bringing them into the 0-10 range. If the sum of the points is less than -350,000, then the user has indicated that none of these drugs are available. Inform them that we cannot do much if this is in fact true.

Assume that the infection is severe, as the physician has taken the time to consult our system. Do check to see if the patient can take drugs orally.

Now, to consider the area in which the infection was acquired:

- If the person is in Thailand or Vietnam: Resistances to chloroquine, Fansidar, and mefloquine (a drug in the same family as quinine) are already a problem. Here, prescribe a joint course of quinine and an antibiotic such as tetracycline

or doxycycline. Fansidar and chloroquine lose 10000. Quinine gains +10000, artemether gains 20000, and the antibiotics will be taken into account later.

- If the person is anywhere else in Southeast Asia, in China, or in Afghanistan/Pakistan, Central America east of the Panama Canal, South America except for Argentina/Chile, or Sub-Saharan Africa: Assume resistance to chloroquine and Fansidar. Take 10000 points away from Fansidar and chloroquine. Add 10000 to quinine.

- If the person is from the Middle East (except Turkey, Syria, or Iraq), Saharan Africa, or a tropical area I have managed to miss: Assume chloroquine resistance only. Add 10000 to quinine and take 10000 points from chloroquine.

- If the person is from anywhere else: There are no malarial drug resistances as of 1992, the time of the cited paper, the most recent source of consolidated data to our knowledge. While individual pieces of data more recent than this are known, and in fact this field evolves as the parasite evolves, the data needs to be frozen at some point. Add 5000 to quinine and 10000 to chloroquine, as chloroquine works on severe infections here.

Other factors:

If the patient is epileptic, or has psoriasis, deduct 6 points from chloroquine's score. This was the interpretation of "some caution"; it moves chloroquine below artemether, but still considers it as a viable option.

If the patient is a pregnant woman, subtract 10000 points from Fansidar, as it should only be given to pregnant women when there is no other alternative.

If the patient is anemic: the score of Fansidar is set to the score of chloroquine, plus 1. the score of chloroquine is set to the score of Fansidar (before the immediately previous change), minus 1. This is an interpretation of "replace chloroquine with Fansidar", which is the recommendation for anemic patients.

Now the drug with the most points is chosen.

A.2 Prescription of Selected Drug

A.2.1 Quinine

All of these numbers are for the salt quinine dihydrochloride.

If unable to take drugs orally:

If the patient is under 13: Prescribe a loading dose of 15mg*weight in kg on day 1, followed by 10mg*kg every 12 hours. Administer intravenously, each dose diluted and given over 2 hours. Same intravenous. Over 13: initial loading dose is 20mg*weight, followed by 10mg*kg every 8-12 hours. If the patient already has quinine in their blood: make the loading dose equal to the maintenance doses. For all: switch to giving medication orally as soon as the patient can handle oral medication. If more than 48 hours of parenteral treatment is expected, reduce maintenance dose by one-half to one-third, unless the patient has renal or hepatic impairment.

If they can swallow: 10mg*kg, 3 times a day, for up to 10 days.

If in Thailand or Vietnam: Same as above - but combine the 7 day quinine course with 250 mg. tetracycline 4 times a day, or 100 milligrams of doxycycline once per day. If this is not possible, or if you are in northwestern Vietnam (where this may not be effective), give a 3-5 day course of quinine, followed by a single dose of Fansidar or a course of antibiotic.

A.2.2 Chloroquine

If patient cannot swallow: Give 25 mg*kg continuously intravenously over the next 24 hours. (or 5 mg*kg 6-hourly) Be very very careful doing this for small children; here, under 8 was chosen as a cutoff.

Otherwise: Give 10 mg*kg on days 1 and 2, and 5 mg*kg on day 3. Otherwise, give 10 mg*kg at start of treatment, and 5 mg*kg 6-8 hours later and on days 2 and 3.

A.2.3 Fansidar

If the patient is under 5, give half a tablet. Each tablet consists of 500 mg sulfadoxine and 25 mg pyrimethamine. If they are between 5 and 9, give 1 tablet; between 9 and 15, 2 tablets; above 15, 3 tablets.

A.2.4 Artemether

Give 3.2 mg*kg on day 1, and 1.6 mg*kg days 2-7. Artemether is dissolved in peanut oil and given by intramuscular injection.

Appendix B

Drug Table

What percentage of this drug is absorbed?
Quinine: 80
Chloroquine: 70-75

When does the peak concentration occur?
Quinine: 1-3 hours after the dosage.
Chloroquine: 2 hours after the dosage.

What is the mean time to elimination?
Quinine: 11 hours in a healthy person, 16 in benign malarias, 18 in severe ones.
Chloroquine: 6-10 days.
Fansidar: 5-8 days for the sulfadoxine portion; 2 weeks for pyrimethamine.

Have any resistances to it developed? Where?
Quinine: Yes, in Thailand and Vietnam. However, sometimes it can be combined with tetracycline or Fansidar in these areas and still be effective.
Chloroquine: Yes, in nearly all malarious areas. It can only be used for severe cases in Central America west of the Panama Canal, Haiti, the Dominican Republic, and parts of the Middle East and West Africa.
Fansidar: Yes, in southeast Asia, Indonesia, and the Amazon.
Artemether: None known. It has mainly been used in China thus far.

How is it administered?
Quinine: Orally, or through intramuscular injection into the inner thigh. Intravenous administration is safe if diluted and given over 4 hours.
Chloroquine: Orally. Additionally, intramuscular and subcutaneous doses can be given; give small doses frequently (3.5 mg/kg base 6-hourly or 2.5 mg/kg base 4-hourly.)
Fansidar: Oral tablets. The intramuscular formulation clears parasitemia as fast as

chloroquine, but produces a less dramatic clinical cure.
Artemether: It is dissolved in peanut oil and given by intramuscular injection.

What are the typical doses administered?
Quinine: For severe malaria, in adults: a 20 mg/kg loading dose (less if they have received quinine in the past 8-12 hours) followed by 10 mg/kg every 8-12 hours.
Severe malaria, in children: 15 mg/kg loading dose, followed by 10 mg/kg every 8-12 hours.
In Vietnam or Thailand: give a 7 day course of quinine in conjunction with tetracycline (250 mg. 4 times a day, for 7 days.) In northwest Vietnam this is no longer effective; however, a short (3-5 day) course of quinine followed by a single dose of Fansidar or a course of tetracycline can be effective.
For normal malarias, it is normally not given due to cinchonism; however, you can administer 10 mg/kg 3 times a day for up to 10 days.
Quinine is often a salt; these numbers are for quinine dihydrochloride. Convert accordingly...
Chloroquine: For normal malarias: 10 mg/kg on days 1 and 2, and 5 mg/kg on day 3. Alternatively, 10 mg/kg initially, 5 mg/kg 6-8 hours later and on days 2 and 3.
Severe malarias: The normal dose can be given as a continuous intravenous infusion over 30 hours. Even better, 10 mg/kg initially over 8 hours, then 15 mg/kg base 8-hourly over the next 24 hours.
Be very careful giving chloroquine non-orally to small children.
Fansidar: Each tablet is 500 mg of sulfadoxine and 25 mg of pyrimethamine. Ages 15 and up, administer 3 tablets; 2 for ages 9-15; 1 for ages 5-9; and 1/2 for children under 5.
Artemether: A typical dose is 200 milligrams the first day and 100 mg on days 2 through 7. Broken down by weight, this becomes 3.2 mg/kg on day 1 and 1.6 mg/kg on days 2-7.

Is it used for normal malarias?
Quinine: Usually not, due to the side effects. However, it can be.
Chloroquine: It is the drug of choice.
Fansidar: Yes.
Artemether: It is good for multiresistant malarias.

Is it used for severe malarias?
Quinine: It is the drug of choice, especially with the advent of widespread chloroquine resistance.
Chloroquine: It works very well in the few cases where the malaria is susceptible. (However, if it is not known that the malaria will be susceptible, it is typically better to go straight to quinine.)
Fansidar: It is better to use one of the other drugs if it is available and will be effective.
Artemether: It's good for multiresistant malarias.

What are its side effects?
Quinine: Cinchonism (temporary deafness, ringing in the ears, headache, dizziness, rash) is bad enough that unsupervised patients will often not complete long courses of quinine. The most common severe side effect is hyperinsulaemic hypoglycaemia. Overdose can be a problem in healthy patients, but is rare in malarious ones.
Chloroquine: Unless the concentration is allowed to get too high, it is generally well tolerated.
Fansidar: They certainly exist.
Artemether: None have been observed in humans; some have been observed in animal studies, which are comparing its efficacy to quinine.

What conditions in the patient should cause caution?
Quinine: none known. It is typically only given in severe cases, however, due to the side effects.
Chloroquine: Epilepsy and psoriasis. If the patient has anemia, chloroquine should be replaced with Fansidar.
Fansidar: Pregnancy.
Artemether: None known.

Can this drug be given to pregnant or lactating women?
Quinine: Conventional doses are safe in pregnancy, even in the final trimester.
Chloroquine: Yes.
Fansidar: Not unless no other drug is available.
Artemether: Yes.

B.1 Other Drug Factors

Other things to consider, no matter the drug:

- Some factors in medication prescription: age, genetic origin, presumed immune status, and, if the patient is female, whether or not they are pregnant or lactating.
- Calculate dosages based on patient's weight when possible.
- If the patient is suffering from severe falciparum, or vomiting or otherwise unable to retain medicine, administer the prescribed dose parenterally. (The normal treatment is to immediately set up a quinine IV, typically quinine dihydrochloride.)
- When malaria is severe, more severe side effects (like cinchonism), that would not be acceptable in the treatment of normal malarias, become more acceptable.
- Consider prior treatment. If a person already has medication in their blood, a loading dose to an unmedicated person may prove toxic. Since drugs such

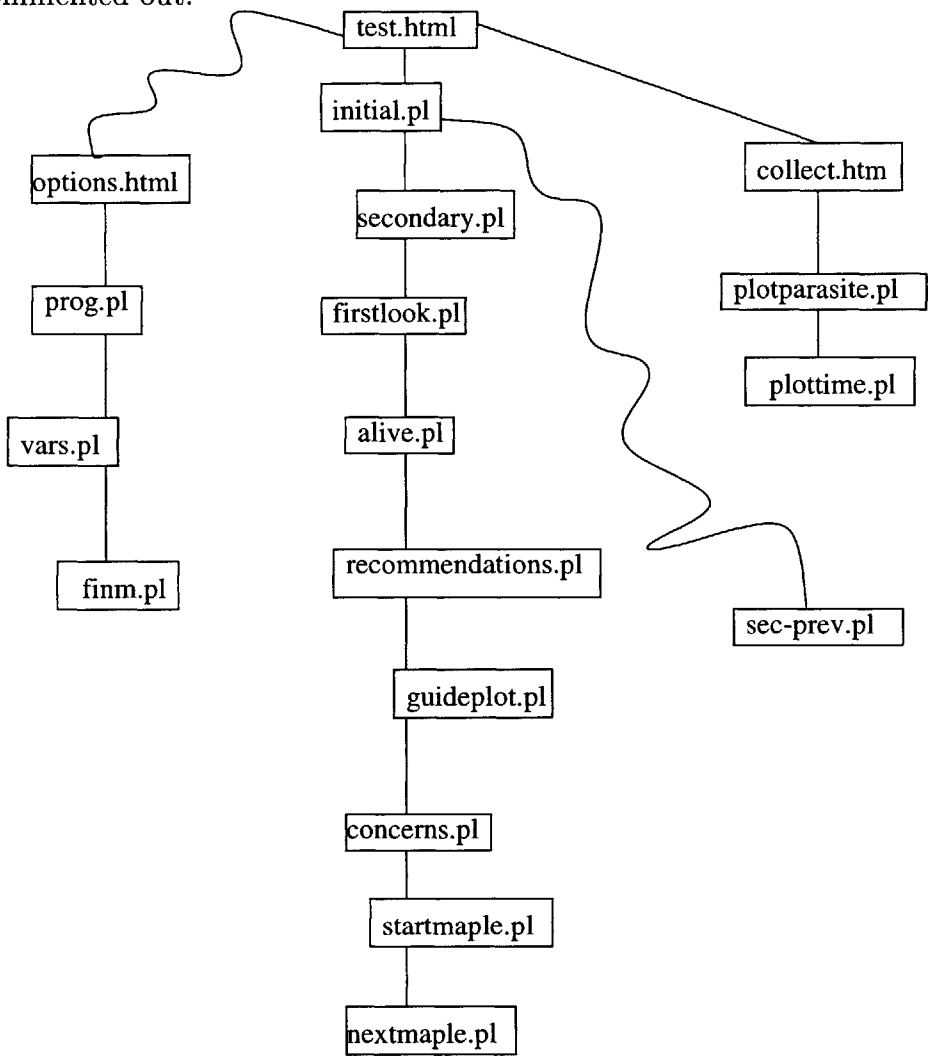
as chloroquine can be obtained over the counter for the home treatment of presumed malaria, this is quite possible.

- Cost is a factor in administering drugs, especially in impoverished areas. On a scale where chloroquine costs 1.0, Fansidar costs 1.6 and quinine costs 7.5.
- Although the treatment of malaria is predominantly done with single drugs, another possibility to consider is drug combinations. This is sometimes done with quinine, especially with antibiotics like tetracycline.
- In addition to the malaria itself, there are side effects of the malaria that have to be fixed, like rehydration, reducing fever, anemia. If the patient has severe malaria, check blood sugar frequently, as hypoglycemia is the most common side effect.

Appendix C

File Hierarchy

This diagram shows the file structure of the program. The curved lines are files which are currently commented out.



Appendix D

Data Recording

D.1 Initial Consultation

It stores its data, in comma-delineated fields, in the file firstname_lastname_main. This data is in the form:

lastname, firstname, age (as a decimal figure in years), weight (can enter in pounds - is stored in kilograms), Celsius temperature (temp. can be entered in Fahrenheit in this section; the program immediately converts it), parasite count per cubic milliliter of blood, time in seconds since epoch (January 1 1970 12:00:00am), whether they cannot take drugs orally, whether they have used the system previously (this is currently disabled leaving a blank space), their gender, whether they are pregnant, if they have anemia, if they have epilepsy, if they have psoriasis, what area they are in, where they are if they said "other" (otherwise blank space), the available drugs with hyphens separating them, the chosen drug, the recommended dosage regimen. Any brackets correspond to where HTML tags were displayed to the user.

This is the file generated for the example patient in the thesis.

Krivan,Mark,23.08,68,37.2,10,915148860,no,,male,,no,no,no,Other,United States,quinine-chloroquine-Fansidar-artemether,chloroquine,Give the patient 680 milligrams of chloroquine base on the first and second days and 340 milligrams on the third; for a total of 1700 milligrams.<p>

D.2 Data Collection

The data is stored in a database tagged with either the nodrug or drug suffix (depending on which case the patient is in). The information is stored in the form of time(seconds since epoch), temp (Celsius), parasite count, drug - dosage, any other recommendations.

In the event there is no drug, it says "no medication" in the drug-dosage field. If a line is instead taken from the person's main file, it puts in the drug-dosage regimen recommended in that file, and then says "first visit" in the comments field.

The form of the database is different from that of the main file, as less information is collected. The data is stored in sorted order (the sort is done with the standard perl command sort) in a file with an ending of nodrug.txt or drug.txt. If the patient has not been given any drug therapy at all, the file is tagged with the nodrug ending, so as to allow those collecting the data to know this fact. If the patient has undergone drug therapy for malaria at any time that the system knows of, the tag is switched to _drug.

This is the example patient's file, after all of the points have been entered.

915408060,39.7,10000,chloroquine-Give the patient 680 milligrams of chloroquine base on the first and second days and 340 milligrams on the third; for a total of 1700 milligrams.<p>,first visit
915494460,37.2,12500,no medication,
915580860,40,20000,totaquine-5 grains,
915645660,37.8,12500,no medication,
915710460,39.7,17500,totaquine-5 grains,
915753660,38.9,22500,no medication,
915840060,39.4,25000,totaquine-5 grains,
915926460,37.8,15000,totaquine-5 grains,
916012860,37.2,17500,no medication,
916099260,37.8,15000,no medication,
916185660,37.8,20000,no medication,

916272060,38.3,15000,no medication,

916358460,37.2,30000,no medication,

916444860,38.3,15000,no medication,

916531260,37.2,20000,no medication,

916617660,40,12500,no medication,

916704060,37.2,15000,no medication,

916790460,38.3,10000,no medication,

916876860,37.2,10000,no medication,

916963260,37.8,10000,no medication,

917049660,37.8,10000,no medication,

917136060,37.2,10000,no medication,

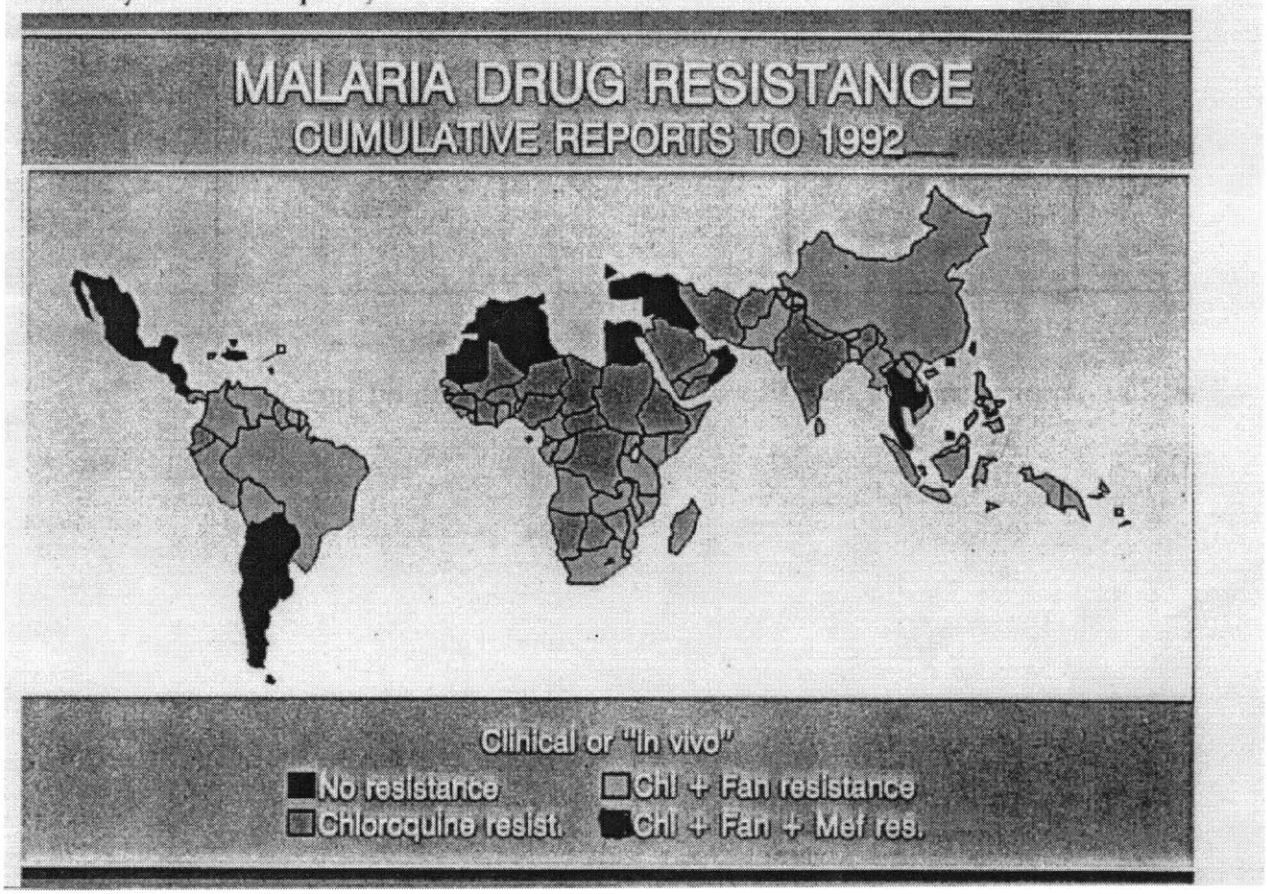
917222460,37.2,10000,no medication,

Appendix E

Malarial Resistances

The following map, taken from *Bruce-Chwatt's Essential Malariology*[5], shows the countries in which malaria is currently known, and the drug resistances that are known in each, as of 1992. As mefloquine is in the same family of drugs as quinine, resistance to one implies resistance to the other.

Plate 28 Global distribution of multidrug-resistant malaria 1992 (Reproduced by courtesy of Dr Shapira.)



Bibliography

[1] http://www.eas.asu.edu/~drapkin/556/mycin.html.

[2] http://www.who.int/ctd/html/malaria.html.

[3] http://www.itknowledge.com/tpj/whatisperl.html.

[4] Mark F. Boyd. *Malariology; A Comprehensive Survey of All Aspects of this Group of Diseases from a Global Standpoint.* W. B. Saunders, Philadelphia, 1949.

[5] H.M. Gilles and D.A. Warrell. *Bruce-Chwatt's Essential Malariology: Third Edition.* Oxford University Press, New York, 1993.

[6] William Long. Temporal reasoning for diagnosis in a causal probabilistic knowledge base. *Artificial Intelligence in Medicine. Vol. 8, pp. 193-215*, 1996.

[7] Howard Silverman. *A Digitalis Therapy Advisor.* Project MAC, Cambridge, 1975.

[8] Ong Lean Suan. Computer-aided diagnosis and treatment of malaria: The imex system. *Comput. Biol. Med. Vol. 20. No. 5, pp. 361-372*, 1990.

[9] William Swartout. *A Digitalis Therapy Advisor With Explanations.* MIT Laboratory for Computer Science, Cambridge, 1977.

[10] Patrick Winston. *Artificial Intelligence, Third Edition.* Addison-Wesley Publishing Company, Reading, MA, 1992.