

1996

## The Diagnosis of Lyme Disease Using an Expert System

Corey May  
*University of Rhode Island*

Follow this and additional works at: <https://digitalcommons.uri.edu/theses>

Terms of Use

All rights reserved under copyright.

---

### Recommended Citation

May, Corey, "The Diagnosis of Lyme Disease Using an Expert System" (1996). *Open Access Master's Theses*. Paper 749.

<https://digitalcommons.uri.edu/theses/749>

This Thesis is brought to you by the University of Rhode Island. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact [digitalcommons-group@uri.edu](mailto:digitalcommons-group@uri.edu). For permission to reuse copyrighted content, contact the author directly.

THE DIAGNOSIS OF LYME DISEASE

USING AN EXPERT SYSTEM

BY

COREY MAY

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

STATISTICS

UNIVERSITY OF RHODE ISLAND

1996

MASTER OF SCIENCE THESIS

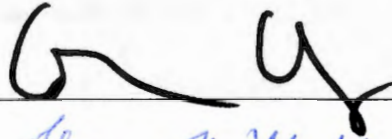
OF

COREY MAY

APPROVED:

Thesis Committee

Major Professor



*Thomas H. Matye*

*James G. Kowalik*

*Thomas J. Rockett*

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

1996

## ABSTRACT

Diagnosing Lyme disease has been problematic since its first recognition in 1975. An assortment of problems, including clinical symptoms that mimic several other diseases, and lack of an accurate laboratory test, have hindered diagnosis. Overdiagnosis and misdiagnosis may result. This thesis seeks to improve the accuracy of diagnosing Lyme disease by creating an expert system.

The type of expert system developed in this thesis will be a probabilistic Bayesian belief network. The network consists of nodes which represent diagnostic variables and links between nodes which represent the probabilistic influence one node has on another. Much is known about Lyme disease, its transmission, and the diagnostic symptoms that are associated with the disease. This information about variables is incorporated into the network through a literature search. Initial estimates of these variables were determined to initialize the system with *a priori* values. To test the system, data were collected on a number of patients who presented symptoms consistent with Lyme disease. The system's classification will be compared to the patients classification based on serological results and methods for improving the system's accuracy are discussed.

## ACKNOWLEDGMENTS

In completing this thesis, I would like to thank Dr. Colleen Kelly for her effort, time, patience and guidance. I would like to thank my committee members Dr. Tom Mather, Dr. James Kolwaski, and Dr. John Montgomery for their time and suggestions. I would also like to thank Joshua Giuliano for the extended use of his laptop. Finally, I would like to thank my parents and sister for their encouragement and support.

## TABLE OF CONTENTS

	Page
Abstract	ii
Acknowledgements	iii
Table of Contents	iv
List of Tables	vi
List of Figures	ix
Chapter 1 Introduction	1
Section 1.1 Bayesian belief expert system	2
Section 1.2 Other classification systems	4
Chapter 2 Bayesian Belief Network	6
Section 2.1 Creating the graph of nodes	6
Section 2.2 Probability propagation in a graph of nodes	7
Section 2.3 Creating the permanent tree of cliques	12
Section 2.4 Probability propagation in a tree of cliques	15
Chapter 3 Initial Estimates	19
Section 3.1 Bayesian Updating	19
Section 3.2 Assumptions made	25
Chapter 4 Methods and results	27
Section 4.1 Methods	27
Section 4.2 Results	29

<b>Section 4.3</b>	<b>Analysis</b>	<b>31</b>
<b>Section 4.4</b>	<b>Assessment of physicians diagnostic ability</b>	<b>34</b>
<b>Chapter 5</b>	<b>Discussion and further research</b>	<b>37</b>
<b>Tables</b>		<b>39</b>
<b>Figures</b>		<b>67</b>
<b>References</b>		<b>75</b>
<b>Bibliography</b>		<b>80</b>

## **LIST OF TABLES**

	Page
Table 1.	Link 1: Initial estimates of the probability of sex and month of onset given Lyme disease.....39
Table 2.	Link 2: Initial estimates of the probability of recalling a tickbite given Lyme disease.....41
Table 3.	Link 3: Initial estimates of the probability of myalgia given Lyme disease.....42
Table 4.	Link 4: Initial estimates of the probability of arthralgia given Lyme disease.....43
Table 5.	Link 5: Initial estimates of the probability of fever given Lyme disease.....44
Table 6.	Link 6: Initial estimates of the probability of malaise or fatigue given Lyme disease.....45
Table 7.	Link 7: Initial estimates of the probability of tender points given Lyme disease.....46
Table 8.	Link 8: Initial estimates of the probability of stiff neck given Lyme disease.....47
Table 9.	Link 9: Initial estimates of the probability of headache given Lyme disease.....48
Table 10.	Link 10: Initial estimates of the probability of sleep disorder given Lyme disease.....49



Table 11.	Link 11: Initial estimates of the probability of meningitis given Lyme disease.....	50
Table 12.	Link 12: Initial estimates of the probability of radiculoneuropathy given Lyme disease.....	51
Table 13.	Link 13: Initial estimates of the probability of encephalitis given Lyme disease.....	52
Table 14.	Link 14: Initial estimates of the probability of Bell's palsy given Lyme disease.....	53
Table 15.	Link 15: Initial estimates of the probability of nodal block given Lyme disease.....	54
Table 16.	Link 16: Initial estimates of the probability of oligoarticular arthritis and age given Lyme disease.....	55
Table 17.	Link 17: Initial estimates of the probability of ECM given Lyme disease.....	56
Table 18.	Link 18: Initial estimates of the probability of serum test given Lyme disease.....	57
Table 19.	Patient data from Wood River Health Services.....	58
Table 20.	Patient data from Dr. Joseph England's office.....	59
Table 21.	Patient data from South County Physicians.....	60
Table 22.	Expert system classifications with and without a serum test.....	61
Table 23.	Results from Fisher's exact test for independence.....	64
Table 24.	A table of the misclassified patients.....	65

Table 25. A table of early and late Lyme patients.....66

Figure 1  
Figure 2  
Figure 3  
Figure 4  
Figure 5  
Figure 6  
Figure 7  
Figure 8  
Figure 9  
Figure 10  
Figure 11  
Figure 12  
Figure 13  
Figure 14  
Figure 15  
Figure 16  
Figure 17  
Figure 18  
Figure 19  
Figure 20  
Figure 21  
Figure 22  
Figure 23  
Figure 24  
Figure 25  
Figure 26  
Figure 27  
Figure 28  
Figure 29  
Figure 30  
Figure 31  
Figure 32  
Figure 33  
Figure 34  
Figure 35  
Figure 36  
Figure 37  
Figure 38  
Figure 39  
Figure 40  
Figure 41  
Figure 42  
Figure 43  
Figure 44  
Figure 45  
Figure 46  
Figure 47  
Figure 48  
Figure 49  
Figure 50  
Figure 51  
Figure 52  
Figure 53  
Figure 54  
Figure 55  
Figure 56  
Figure 57  
Figure 58  
Figure 59  
Figure 60  
Figure 61  
Figure 62  
Figure 63  
Figure 64  
Figure 65  
Figure 66  
Figure 67  
Figure 68  
Figure 69  
Figure 70  
Figure 71  
Figure 72  
Figure 73  
Figure 74  
Figure 75  
Figure 76  
Figure 77  
Figure 78  
Figure 79  
Figure 80  
Figure 81  
Figure 82  
Figure 83  
Figure 84  
Figure 85  
Figure 86  
Figure 87  
Figure 88  
Figure 89  
Figure 90  
Figure 91  
Figure 92  
Figure 93  
Figure 94  
Figure 95  
Figure 96  
Figure 97  
Figure 98  
Figure 99  
Figure 100

## **LIST OF FIGURES**

	Page
Figure 1.	The graph of nodes before triangulation.....67
Figure 2.	The graph of nodes after triangulation.....68
Figure 3.	The tree of cliques.....69
Figure 4.	The partial list of configurations inside a particular clique.....70
Figure 5.	A graph of nodes with the addition of diagnostic historical variables.....71
Figure 6.	A plot of the Beta distribution.....72
Figure 7.	A scatter plot of physician's likelihoods of Lyme disease vs. expert system's probabilities of Lyme without a serum test.....73
Figure 8.	Box plots of physicians likelihoods of Lyme and non-Lyme patients.....74

## **Chapter 1: Introduction.**

Diagnosing Lyme disease has been problematic since the disease's identification in 1975. An assortment of problems, including clinical symptoms that mimic a number of other diseases and lack of an accurate laboratory test, have hindered diagnosis. Despite these problems, much is known about the spatial distribution of the disease, the mechanisms of acquiring the disease, and the progression of the disease.

Lyme disease is a dangerous disease that continues to challenge scientists and clinicians. It is a complex and enigmatic spirochetal illness involving the skin, nervous system, heart, and joints. It is the most common tick-borne disease reported by the Center for Disease Control (CDC) (Magnarelli, 1988). Incidence in the state of Rhode Island has been among the highest in the country: CDC reported incidences of 0.000272 in 1992 and 1993 (MMWR). Effective treatment of Lyme disease requires early diagnosis. If the disease is allowed to progress, dangerous side effects and symptoms may develop, such as meningitis, encephalitis, Bell's palsy, radicular pain, nodal block, and arthritis. With early diagnosis, oral antibiotics such as amoxicillin or doxycycline are usually curative (Barbour, 1993).

Unfortunately, there are difficulties with the diagnosis of Lyme disease. Like syphilis, Lyme disease has been called a "great imitator" and is often difficult to diagnose clinically, particularly when a skin lesion, erythema chronicum migrans (ECM), is absent. This may occur in 10-40% of patients (Dennis, 1991). The presence of ECM definitely implies the patient has Lyme disease, however, many physicians may not accurately diagnose ECM. Blaauw (1987) tested the diagnostic ability of general practitioners and dermatologists using photographs of ECM and found their recognition lacking.

Laboratory tests are not yet standardized and scientists have highlighted the poor agreement of test results using serological methods. On a national level, the percentage of false negative test results are estimated to be 4% to 21% and the percentage of false

positive rates are estimated to be 2% to 7% (Kahneman, 1982). False negative tests may lead to serious problems if the disease is allowed to progress untreated; false positive tests may lead to complications by overtreating with antibiotics.

The rate of false positives and the inability of most serological tests to distinguish between active and inactive disease has led to the overdiagnosis of Lyme in some areas. Lyme disease is increasingly reported from the southeastern and south-central regions of the United States, even though the causative agent, *Borrelia burgdorferi*, has not been found in most states in these regions (Dennis, 1991). Fibromyalgia and Chronic Fatigue Syndrome (CFS) may be incorrectly diagnosed as advanced Lyme disease, due to the similarity of their clinical symptoms. The problems in distinguishing these diseases are compounded by the fact that some patients develop Fibromyalgia or CFS in association with or soon after Lyme disease. For example, Steere (1993) found that 45% of the patients diagnosed with a disease other than Lyme also previously tested positive by serological assay for *Borrelia burgdorferi*, a fact which may reflect the positive association of these diseases.

Diagnosing Lyme disease can be difficult even for the physician who is well-informed about the clinical manifestations of the disease. The diagnosis problem involves an assessment of a number of clinical symptoms, possible risk factors and the timing of the symptoms. Cognitive psychologists have shown that the ability of the human brain to make these types of assessments of uncertainty in complicated situations is generally poor (Kahneman, 1982). This thesis seeks to improve diagnosis of Lyme disease by developing a computerized expert system.

### **1.1 Bayesian Belief Expert System:**

An expert system is a computer program which can make reasonable judgments in a complex area (Neapolitan, 1990). The intent is that the program has the capability to perform a task ordinarily performed by an expert. Thus even physicians unfamiliar with

Lyme disease could diagnose with the accuracy of an expert with this tool. Not only could such an expert system improve diagnostic accuracy, but it could also save time and money for health care providers. Examples of wasted costs are laboratory tests, physicians' services, hospital care, and unnecessary medications (Magid, 1992). Thus, an expert system that diagnoses Lyme disease with a high certainty would represent a significant advancement.

In diseases such as Lyme disease a probabilistic expert system is more appropriate than a deterministic system. Because many of the clinical symptoms of Lyme disease are also symptoms of other diseases, the presence of these symptoms increases the probability of having Lyme disease, without implying that the patient definitely has Lyme disease. In a deterministic system, a diagnostic variable directly implies that a particular state of the disease variable has occurred.

The paradigm used to build the expert system in this thesis is a Bayesian Belief network. A Bayesian belief network is a tree-like structure containing a number of diagnostic variables, referred to as nodes. Links between nodes represent the probabilistic influence of one node on another. The tree-like structure is modeled so that nodes are connected by the links in a parent/child arrangement according to causality. Each link carries probabilities which measure the influence that each parent has on its connected child. One objective of this thesis was to obtain estimates of these links by searching the existing literature, and by obtaining opinions of expert Lyme diagnosticians. Once these estimates have been obtained, information on a given patient's symptoms can be entered in the system, and the system will calculate the probability that the patient has Lyme disease. The Bayesian belief network has three engaging features: it easily incorporates information from a variety of sources, it may be updated with new data and it handles missing data very well.

## 1.2 Other classification systems:

Three other probabilistic classification methods were considered for use in this project: discriminant analysis, a log-linear model, and a classification and regression tree. These methods were ultimately decided against, but are briefly described here for completeness. Discriminant analysis is a classification method using continuous, normally distributed explanatory variables. Since most explanatory variables in this study are categorical, a discriminant analysis was not chosen. Log linear models could be used for classification problems with both categorical and continuous explanatory variables. A classification and regression tree splits patients into disease categories (e.g. early, late, or no Lyme disease) based on the values of the diagnostic variables. However, estimation in these models is difficult or inefficient with missing data, and interaction terms don't fully represent a cause and effect relationship between explanatory variables. In addition, the ability to combine information from a variety of sources, and to learn from new data are not possible with these classical statistical methods.

This thesis develops an expert system that will aid physicians in the diagnosis of Lyme disease. Chapter Two describes in detail how the Bayesian belief network works. Chapter Three discusses how the initial estimates were obtained through a literature search and with the aid of expert's opinions. Chapter Four describes the use of this system on a set of patients and compares a patient's diagnosis as determined by the expert system to the diagnosis determined by serology. Finally, Chapter Five discusses the results and directions for further research.

## **Chapter 2: Bayesian belief network.**

Determining the probability that a patient has Lyme disease given various clinical variables, is the major goal of this thesis. A Bayesian belief network that represents the causal relationship between clinical variables and Lyme disease facilitates calculation of this probability. The initial step in creating a Bayesian belief network involves construction of a graph of nodes which graphically represents the causal relationships among all variables. Probability calculations become difficult in complex networks, therefore, a different representation of this information, a tree of cliques, is used instead. In this chapter, the four steps to create and use a Bayesian belief network are described in detail: (1) creating the graph of nodes, (2) probability propagation in a graph of nodes, (3) creating the permanent tree of cliques, and (4) probability propagation in the cliques.

### **2.1 Creating the Graph of Nodes:**

A Bayesian belief network consists of a number of diagnostic variables represented as nodes. Causal relationships between variables are represented by links; one node points to another if it directly affects the probability of that node occurring. This establishes parent/child relationships where a parental node points to its child node. Figure (1) depicts the cause and effect relationships for the variables used in this study. For example, the arrow pointing from *Age* to *Oligoarticular Arthritis*, indicates that *Oligoarticular Arthritis* could be caused by a person's age. The link from *Disease* to *Serum Test* indicates *Disease* causes a positive or negative serum test, that is, *Disease* influences the



probability of the serum test. It should be noted that this causation is not deterministic but probabilistic. In these examples, the nodes *Age* and *Disease* are the parents of the child nodes *Oligoarticular Arthritis* and *Serum Test*, respectively. A node may have more than one parent; in Figure 1 *Oligoarticular Arthritis* has two parents, *Disease* and *Age*. This means that *Oligoarticular Arthritis* could be caused by *Age* or *Disease* or both.

Each link in the graph of nodes has an associated probability matrix, **P**, that specifies the amount of influence that the parent node has on the child. The matrix has dimensions (N x M) where N is the number of states in the parent node, and M is the number of states in the child node, and has elements  $P_{ij} = P(\text{child} = c_j \mid \text{parent} = s_i)$  where  $c_j, s_i$  are the possible child states and parent states, respectively,  $i = 1, 2, \dots, N$ , and  $j = 1, 2, \dots, M$ . For example, in Figure 1,  $P_9$  links the parent node *Disease* to the child node *headache*. If the states of Lyme disease are defined as  $L_0$  = no Lyme disease;  $L_1$  = early Lyme disease;  $L_2$  = late Lyme disease; and  $L_3$  = Fibromyalgia or Chronic Fatigue Syndrome, and the child states are defined as  $H_1$  = has a headache and  $H_2$  = doesn't have a headache, then the probability matrix for this example is:

$$P_9 = \begin{bmatrix} P(H_1 \mid L_0) & P(H_2 \mid L_0) \\ \cdot & \cdot \\ \cdot & \cdot \\ P(H_1 \mid L_3) & P(H_2 \mid L_3) \end{bmatrix}$$

The first column of the matrix contains the probabilities of having a headache given each state of the parent node (Lyme disease) and the second column of the matrix would have the probability of not having a headache given each state of the parent node.

Each node also has an associated belief vector whose elements are the marginal probabilities of each state of that node. For example, the belief vector for the Lyme disease node is  $(4 \times 1)$  and has elements equal to the incidences of each state of Lyme disease. We will define the belief vector of the Lyme disease node as follows:

$$B(L) = \begin{bmatrix} P(L_0) \\ P(L_1) \\ P(L_2) \\ P(L_3) \end{bmatrix}$$

## 2.2 Probability propagation in a Graph of nodes:

Calculating the probability of Lyme disease given a set of diagnostic variables is accomplished by sending lambda and pi messages. Suppose that  $E$  represents the set of observed diagnostic variables;  $E^+_A$  are the variables observed in the tree above node  $A$  and  $E^-_A$  are the variables observed at or below node  $A$ . Lambda messages send information about  $E^-_A$  up to node  $A$ ; pi messages send information about  $E^+_A$  down to node  $A$ .

Each node  $A$  contains a lambda value defined as:

$$\Lambda(A) = P(E^-_A | A)$$

If  $A=a_j$  is observed, then this lambda value is a vector of zeroes with a one in position  $j$  where the vector has dimension  $(K \times 1)$  where  $K$  is the number of states in node  $A$ .

Otherwise this value must be computed recursively via lambda messages. For example, if

a patient had a headache, yes in the node *Headache*, would be instantiated. The lambda value for this node is  $\Lambda(\text{headache}) = (1, 0)$ . Likewise, instantiating for not having a headache results in a the lambda value of  $(0, 1)$ . All lambda values initially start at  $(1, 1)$  for each node, which indicates an absence of any diagnostic evidence. After a node has been instantiated the lambda values change. Note that  $\sum_i \Lambda(A_i)$  is not necessarily one.

Information is passed upward in the tree through lambda messages in the following way. The lambda message from child node B to parent node A is defined as:

$$\Lambda_B(A) = P(E^-_B | A).$$

The lambda message is a vector with dimension  $(K \times 1)$  where K is the number of states in parental node A. A lambda message from B to A thus sends information to A from the variables below B. Lambda messages can be calculated as:

$$\Lambda_B(A) = \sum_i P(B_i | A) * P(E^-_B | B_i) = P * \Lambda(B);$$

the first equality is a consequence of the law of total probability and the conditional dependence implied by the graph structure.

As an example, suppose a patient reported experiencing a headache. Then the lambda message from the node *headache* to the *Lyme Disease* node would be computed as follows:

$$\begin{bmatrix} P(H_1 | L_0) & P(H_2 | L_0) \\ \cdot & \cdot \\ \cdot & \cdot \\ P(H_1 | L_3) & P(H_2 | L_3) \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} P(H_1 | L_0) \\ P(H_1 | L_1) \\ P(H_1 | L_2) \\ P(H_1 | L_3) \end{bmatrix}$$

The lambda message is the probability matrix for the link times the lambda value of the child. After all lambda messages are received, the parent's lambda value is updated by multiplying the lambda messages of all its children. This is done because all of node A's children are independent of each other given A. If  $s(A)$  is the set of A's children, then

$$\Lambda(A) = \prod_{c \in s(A)} \Lambda_C(A),$$

where the product indicates elementwise multiplication.

Pi messages are sent down the tree, relaying direct information from the parents to the children. Each node has a pi value defined as:

$$\pi(A) = P(A | E^+_A)$$

When A has no parents, A is a root node, and its pi value is equal to  $P(A | \emptyset) = P(A)$ .

The pi message from node A to B is defined as:

$$\pi_A(B) = B'(A) / \Lambda_B(A)$$

where  $B'(A)$  is the current belief in node A and the division is elementwise. The pi message is a vector with dimension  $(K \times 1)$  where K is the number of states in parental node A. The belief vector is divided by the lambda message from node B to A so that, the

pi message from node A to B sends information only about  $E^+_A$ . After the pi message is computed it is normalized so the elements in the message sum to one. Information from a parent is incorporated into the child node by updating the child's pi value:

$$\pi(B) = P^T * \pi_A(B) = \sum_i P(B | A_i)P(A_i)$$

This equation again is a consequence of the law of total probability.

For example, suppose the Lyme disease node is instantiated for  $L_0$ , and Lyme disease is the parent of headache. The pi value of Lyme disease node then is equal to (1, 0, 0, 0), by definition. The pi message that is sent from Lyme disease to headache is,

$$\pi_L(H) = (1, 0, 0, 0) / \{ P(H_1 | L_0), P(H_1 | L_1), P(H_1 | L_2), P(H_1 | L_3) \}.$$

After normalization the pi message is (1, 0, 0, 0). Finally, the pi value for the headache node is computed by multiplying the transpose of the link matrix by the pi message that was just sent:

$$\pi(H) = \begin{bmatrix} P(H_1 | L_0) & P(H_1 | L_1) & P(H_1 | L_2) & P(H_1 | L_3) \\ P(H_2 | L_0) & P(H_2 | L_1) & P(H_2 | L_2) & P(H_2 | L_3) \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} P(H_1 | L_0) \\ P(H_2 | L_0) \end{bmatrix}$$

Thus, the pi value for the child is the probability of the node given  $E^+_A$  or the instantiated variable.

After the lambda and pi messages are sent, each node's belief is updated using Bayes theorem:

$$B'(L) = \alpha * \Lambda(L) * \pi(L)$$

where  $B'(L)$  is the new belief of the Lyme disease node due to the new diagnostic evidence,  $\alpha$  is a normalizing constant, and the multiplication is elementwise.

The network is first initialized to compute *a priori* probabilities of all nodes (i.e. the probabilities based on the instantiation of no nodes). The network is initialized through probability propagation. All lambda messages and lambda values are initially set to one. A pi message from the top of the graph is sent down the graph to all of its children whereupon a propagational flow will begin. Each child then sends pi messages to each of its children, and this process is repeated until the leaves of the tree are reached. At this point, each leaf of the tree sends lambda messages up to its parents, which then send lambda messages to their parents. This process is repeated until the root node is reached.

After initialization, the system is ready to calculate the probability of Lyme disease given a set of diagnostic variables. First each diagnostic variable is instantiated; this instantiation will change a number of nodes pi and lambda values. Then this information is propagated throughout the network via lambda and pi messages. Each instantiated node sends a lambda message to its parent and a pi message to its child. If the parent or child node doesn't exist then no message is sent. Once every node has sent both its pi and lambda messages, the propagation is complete.

This method of probability propagation is adequate if no node has more than one parent, and there is no more than one path connecting each pair of nodes. However, this may not be true in more complex networks. If a node has more than one parent, a modification of the probability link matrix may be made that conditions on all parental nodes. With this modification, a single pi message is sent from all parents to the child.

However, if there is more than one path between a pair of nodes, a lambda message from A to B and a pi message from B to A could be sent down both paths, causing errors in the probability propagation method described earlier. A different representation of the graph is necessary to correct these problems.

The Bayesian belief network used here will propagate probabilities in a tree of cliques. Probability propagation could have been done in either a graph of nodes or a tree of cliques because there is not more than one path between any two nodes. The package Dexpress computed the probabilities in a tree of cliques. The next section describes how cliques are determined and how probabilities are propagated in a tree of cliques.

### **2.3 Creating the permanent tree of cliques:**

Once the tree of nodes has been constructed we re-represent this graph as a tree of cliques. As discussed earlier, the tree of cliques allows for easier and more accurate calculations in more complex graphs. A clique is a subset of nodes in the network that is complete and maximal. A set is complete if every pair of distinct nodes in the set are connected with a link. A set is maximal if it is not a subset of any other complete set (Neapolitan, 1990). This prevents counting a clique as two cliques when, in fact, it should be counted as one. The tree of cliques is constructed in three steps: (1) the graph is triangulated, (2) the cliques are determined, (3) and the tree of cliques is determined. The tree of cliques will contain all the information stored in the original graph.

In order to determine the different cliques of the graph, the system must be triangulated. The system is triangulated if every simple cycle of length strictly greater than

three possesses a chord. A simple cycle is a path from one node, back to itself, where no other node on this path can be repeated. A simple cycle possesses a chord if there is a link between two nonconsecutive nodes of the cycle (Neapolitan, 1990).

Triangulation of the system is accomplished in three steps: moralizing the graph, numbering the nodes, and determining the fill-in's. First, the graph is moralized by marrying or joining the parents of a node with a link. Second, the nodes are numbered using an algorithm called Maximum Cardinality Search (MCS) (Neapolitan, 1990). Maximum Cardinality Search runs by assigning the number one to an arbitrary node. Successive nodes are selected as the node adjacent to the largest number of previously numbered nodes, breaking ties arbitrarily. The final step before determining the cliques is to determine the fill-in's or links that must be added to triangulate the graph. Fill-in's occur when there is a path between two nodes  $(v,w)$  containing only  $v$ ,  $w$  and vertices ordered after both  $v$  and  $w$ . If this occurs, then  $v$  and  $w$  must be connected resulting in what is called a fill-in. After completion of these steps, the tree of variables is triangulated (see Figure 2).

Next the cliques are determined. Using the triangulated tree of nodes, the complete and maximal sets of nodes are determined. For example, referring to Figure 1, the nodes *Disease*, *Age*, and *ECM* are a clique because each node is linked to every other node in the set and no additional nodes may be added and still retain the complete property. The total number of cliques for this network is eighteen.

The cliques contain all information of the original graph. All possible combinations of states for all the nodes in a clique, or the configurations, are stored in the clique. In



Figures 1 and 3, clique 3 consists of nodes C and E where node C has four possible states (c1, c2, c3, c4) and node E has two possible states (e1, e2), so the configurations are: (c1, e1), (c1, e2), (c2, e1), (c2, e2), (c3, e1), (c3, e2), (c4, e1), (c4, e2). Information about the causal relationships among the nodes are contained in the sets  $R_i$  and  $S_i$  defined for each clique  $i$ .  $S_i$  is defined as the intersection of clique  $i$  with the union of all previously numbered cliques;  $R_i$  is equal to  $Clq_i - S_i$ , or all other nodes in that clique. The set  $R_i$  for clique 3 contains the node (E). The set  $S_i$  for clique 3 contains the node (C). Psi values and posterior probabilities are also contained in each clique and represent the information previously stored in the link probability matrices and the belief vectors in the tree of nodes, respectively. These values are explained in detail below. Figure 4 has a partial table of the configurations and the previously mentioned sets that must be stored at each clique. Cliques one and two have a large number of configurations, therefore, they are only partially listed but cliques three and four are fully listed.

The tree of cliques is determined in the following way. Remembering that a clique is defined as a set of nodes, we number the cliques by the highest labeled node in the clique, breaking ties arbitrarily. By using the MCS algorithm, the nodes possess the running intersection property (Neapolitan, 1990) defined as: each clique, except the first, when intersected with the union of all the previous cliques, results in a set of nodes that is contained entirely in a previous clique. The placement of cliques into a tree (Figure 3) can now be determined because there is a development of a cause and effect (parent/child) relationship due to the running intersection property. After the cliques are placed into a

tree, the original graph can be disregarded. The tree becomes a permanent part of the expert system and is only changed when the causal network changes

#### 2.4 Probability Propagation in a tree of cliques:

The Bayesian belief network affects diagnoses by calculating the probability of Lyme disease given observation of certain diagnostic variables. As discussed earlier, the graph of nodes propagated probabilities via lambda and pi messages. In a tree of cliques probability propagation is similar. Information that was stored in the link probability matrices is stored instead in a vector of psi values. Belief vectors are defined for configurations rather than for the states of the nodes. Probability propagation once again takes place by sending messages up through the tree of cliques and back down.

The psi value for  $Clq_i$  is a  $N \times 1$  vector defined as,

$$\Psi(Clq_i) = P(R_i | S_i).$$

where  $R_i$ ,  $S_i$  are the sets defined earlier and  $N$  is the number of configurations in clique  $i$ .

The psi values for  $Clq_i$  is calculated as follows:

$$\Psi(Clq_i) = \prod_{f(v)=Clq_i} P(v | c(v)) = P(R_i | S_i).$$

where  $P(v | c(v))$  is the conditional probability of a node ( $v$ ) given its parent ( $c(v)$ ), and  $f(v)$  is the set of all nodes assigned to clique  $i$ . If no node ( $v$ ) is assigned to  $Clq_i$  the product equals 1.

Probability propagation in the tree of cliques is done in two steps. First, lambda messages from the leaves (nodes with no children), or the bottom of the tree, propagate information up the tree. Which leaf starts the propagation and the order in which the

leaves send their messages are arbitrary. Nodes that have received lambda messages send lambda messages to their parents. The lambda messages propagate upward until the root is reached. Second, after all the lambda messages reach the root of the tree, pi messages are propagated down the tree. Once all the leaves have received pi messages, the propagation is complete.

Instantiation of a variable in a tree of cliques is similar to instantiation in a graph of nodes. Every clique containing a variable that has been instantiated for, multiplies its psi values by a vector of zeroes with ones in the positions corresponding to configurations that have occurred.

Now a lambda message can be sent from a leaf clique to a parent clique and the propagation may begin. The lambda message from  $Clq_i$  to its parent clique,  $\lambda_{clq_i}(S_j)$ , is defined similarly to before:

$$\lambda_{clq_i}(S_j) = \sum_{R_i} \Psi(Clq_i) = \sum_{R_i} P(R_i | S_j) P(E | R_i)$$

where the summation is over all the nodes in  $R_i$  or the nodes that are not contained in the parent clique. The lambda message contains information solely about the nodes in both parent and child cliques or nodes in  $S_j$ , thus the summation over the set  $R_i$ . If  $R_i$  is equal to the empty set, there is no summation. The lambda message is a  $K \times 1$  vector where  $K$  is the product of the number of states of each node in  $S_j$ . If  $S_j$  is equal to the empty set, no message is sent. After computing the lambda message and before leaving the clique, the conditional probabilities are divided by this lambda message,

$$\Psi(Clq_i) = \Psi(Clq_i) / \lambda_{clq_i}(S_j),$$

where the division is elementwise. This normalizes the conditional probabilities so that the marginal probabilities can be calculated. The lambda message that is received by the parent is multiplied elementwise by the parent's current psi values.

The second phase of probability propagation begins at the root of the tree. Once all lambda messages have been sent, the unconditional probabilities,  $P(\text{Clq}_j)$  can be determined. To determine these probabilities we use pi-messages to propagate the information down the tree. A pi-message from  $\text{Clq}_j$  to  $\text{Clq}_i$  is computed as:

$$\pi_{\text{Clq}_j}(\text{S}_i) = \sum_{\text{Clq}_j - \text{S}_i} P(\text{Clq}_j)$$

where the  $P(\text{Clq}_j)$  are the unconditional probabilities for each configuration in the parental clique, and the summation is over all the variables that the parent clique doesn't share with the child clique. The pi message thus pertains only to the variables in the child clique's S set. If  $\text{Clq}_j - \text{S}'_i = \emptyset$  then the sum is  $P(\text{Clq}_j)$ . If  $\text{S}'_i = \emptyset$  the message is 1. The propagation of pi messages down the tree continues until all of the leaves of the tree are reached.

The posterior probabilities in the tree of cliques are similar to the belief vectors from the graph of nodes. The posterior probabilities, or unconditional probabilities, are contained in a vector with the same dimension of its corresponding psi value vector. The posterior probabilities,  $P(\text{Clq}_i)$ , are defined as  $P(\text{R}_i | \text{S}_i)P(\text{S}_i)$ . The posterior probabilities are calculated for each configuration in the tree as follows:

$$P(\text{Clq}_i) = \Psi(\text{Clq}_i) * \pi_{\text{Clq}_j}(\text{S}_i) = P(\text{R}_i | \text{S}_i)P(\text{S}_i)$$

where the multiplication is elementwise.

Finally, the probabilities of each variable in each node can be determined using the law of total probability,

$$P(v) = \sum_{w \in \text{cl}_i, w \neq v} P(\text{Cl}_i | w).$$

This equation sums over all the variables except  $v$  in clique  $i$ . For example, in Figure 3 clique 3 consists of the nodes C and E. The probability of the first state of node C is the summation over all the states of node E with  $c_1$  fixed, or

$$P(c_1) = \sum_i P(c_1, e_i) = P(c_1, e_1) + P(c_1, e_2).$$

The probabilities of other states are found in a similar manner.

### **Chapter 3: Initial estimates.**

In this chapter, we derive estimates of the link probability matrices for the Bayesian belief network through a meta analysis of research papers and collected data. The literature search was conducted using the keywords Lyme disease, ticks, *Ixodes*, Fibromyalgia, and CFS. Journal articles were obtained from the University of Rhode Island and the Brown University libraries. The Centers for Disease Control and Prevention's Lyme disease data base was also analyzed with the help of Dr. Kathy Orloski. Additionally, a survey was administered to physicians to obtain information on symptoms related to patients with diseases other than Lyme disease. This chapter explains how the estimates were obtained, and describes the assumptions made.

#### **Section 3.1 Bayesian Updating:**

The elements of the link probability matrices are the conditional probabilities,  $P(c_i | p_k)$  where  $i = 1, 2, \dots, n$  and  $k = 1, 2, \dots, m$ ,  $n$  being the number of states in the child and  $m$  being the number of states in the parent. Thus it was necessary to estimate the proportion of people reporting each of the symptoms in the network for each disease category. We searched for articles with data on the numbers of patients with various symptoms. Proportions estimated from these data serve as the initial estimates of link probability matrices.

The literature search focused on Lyme disease, CFS and Fibromyalgia, and *Ixodes* ticks. *Ixodes* ticks were included in the search because *Ixodes scapularis* and *Ixodes*

*pacificus* are the species of ticks that transmit Lyme disease. Thirty eight articles were found that had data on patients with early or late Lyme disease, or CFS or Fibromyalgia. Articles were included in the study if they listed the number of patients with various symptoms. Frequencies of these symptoms given the different stages of Lyme disease were recorded.

The Center for Disease Control and Prevention (CDC)'s database contains information on over 30,000 Americans with Lyme disease. The CDC classified a patient as having Lyme disease if the patient had ECM or one or more late manifestations together with positive serology. Estimates for the proportions of meningitis, encephalitis, arthritis, nodal block, Bell's palsy, radicular pain, and onset of disease were obtained from the 29,189 people that met the CDC's case definition and came from an endemic geographical region.

In the literature search only one paper was found with data on symptoms of patients with no Lyme disease: Waylonis and Heck<sup>40</sup> tested a control group against a group of patients with Fibromyalgia. To better estimate the proportions of patients that do not have Lyme disease but have symptoms consistent with Lyme, a survey was distributed to doctors in the South County area of Rhode Island. The survey queried each doctor to the average number of patients without Lyme disease seen in a typical week. These results are incorporated into the system.

The information from the literature search, the CDC database, and the physician surveys are integrated using a Bayesian approach. Bayesians view probabilities subjectively. If there is no information about a particular link element except that its value

must be between zero and one, then this element would be viewed as a random variable with a uniform distribution on the interval ( 0, 1 ). This is the prior distribution. New information about the link would cause the belief in the link element to change and thus the posterior distribution describing the random variable would be updated to reflect the new belief. Updating distributions with the new information is accomplished using Bayes theorem:

$$\text{posterior} \sim \text{prior} * \text{likelihood}$$

where the likelihood is the probability of the data given that the link is known.

We assume that the conditional probabilities in the links have a dirichlet distribution. For nodes with two states, the dirichlet distribution reduces to a beta distribution. The dirichlet distribution is desirable for two reasons. First, the dirichlet and beta distributions model random variables with possible values on the interval (0, 1), therefore, they are appropriate for modeling probabilities. Second, both distributions are conjugate priors, so that they lead to a posterior distribution that is also a beta or dirichlet distribution, respectively. Additionally, updating is easy for both beta and dirichlet distributions.

Initially, each row of the link matrix,  $(P_{i1}, P_{i2}, \dots, P_{ik})$  is modeled with a joint distribution that is uninformative. Spiegelhalter (1993) recommends modeling the joint distribution as:

$$(P_{i1}, P_{i2}, \dots, P_{ik}) \sim \text{dirichlet}(1/k, 1/k, \dots, 1/k)$$

where  $k$  is the number of child states. Then the expected value of the link element  $P_{ij}$  ( the conditional probability that the child state is  $c_j$  given that its parent state is  $s_i$  ) is,



$$E[ P_{ij} ] = E[ P ( c_j | s_i ) ] = 1/k$$

where  $s_i$  is the parent of  $c_j$ , and  $j = 1, 2, \dots, k$ . Thus the probabilities of the child states (given the parent  $s_i$ ) are equally likely, reflecting the fact that nothing is known about the link.

The dirichlet distribution updating method is proved using Bayes' theorem. Bayes' theorem says,

$$\text{posterior} = f(\mathbf{p}_i | \mathbf{x}) \sim f(\mathbf{x} | \mathbf{p}_i) * f(\mathbf{p}_i) = \text{likelihood} * \text{prior}$$

where  $\mathbf{p}_i = (P_{i1}, P_{i2}, \dots, P_{ik})$ , the  $i^{\text{th}}$  row of the link matrix, and  $\mathbf{x} = (x_1, x_2, \dots, x_k)$ . First, suppose the joint prior distribution of  $\mathbf{p}_i$  is Dirichlet(  $a_1, a_2, \dots, a_k$  ), the density function is:

$$f(\mathbf{p}_i) = \binom{n}{a_1-1, \dots, a_k-1} * \left[ (P_{i1})^{a_1-1} (P_{i2})^{a_2-1} \dots (P_{ik})^{a_k-1} \right]$$

where the first term is a constant  $h$ .

Suppose that of  $n$  observed patients with parent  $s_i$ ,  $x_1$  fall into the child category  $c_1$ ,  $x_2$  into the child category  $c_2$  and  $x_k$  into child category  $c_k$ . Suppose that each patient has the same probability of falling into the child categories, i.e.  $(P_{i1}, P_{i2}, \dots, P_{ik})$  and that the patients are independent of one another, then the likelihood or the probability of the data given  $P_{i1}$ , is given by the multinomial distribution:

$$f(\mathbf{x} | \mathbf{p}_i) = \binom{n}{x_1, x_2, \dots, x_k} * \left[ (P_{i1})^{x_1} (P_{i2})^{x_2} \dots (P_{ik})^{x_k} \right]$$

where the first term is a constant  $h'$ . Multiplying the prior and the likelihood,

$$[ h * (P_{i1})^{a_1-1} (P_{i2})^{a_2-1} \dots (P_{ik})^{a_k-1} ] * [ h' * (P_{i1})^{x_1} (P_{i2})^{x_2} \dots (P_{ik})^{x_k} ],$$

gives a quantity proportional to the posterior distribution,

$$h'' * (P_{i1})^{a1+x1-1} (P_{i2})^{a2+x2-1} \dots (P_{ik})^{ak+xk-1},$$

where  $h''$  is a constant. The resulting posterior distribution is dirichlet with parameters  $(a_1+x_1, a_2+x_2, \dots, a_n+x_n)$ . Thus the dirichlet is a conjugate prior and updating is accomplished in the very simple manner of adding cases to the prior parameter values.

This updating procedure will be illustrated using an example. Referring to table 2, link 2 is the probability of recalling a tickbite given you have Lyme disease. Suppose we are updating for the parental state of early Lyme disease and the child states are  $R_1 =$  recall tickbite and  $R_2 =$  don't recall tickbite. Initially the link starts as a vague distribution:

$$(P_{11}, P_{12}) = \{P(R_1 | L_1), P(R_2 | L_1)\} \sim \text{Dirichlet}(0.5, 0.5).$$

Petersen et. al.<sup>16</sup> reported that 417 early Lyme disease patients recalled a tickbite and 362 early Lyme patients didn't recall a tickbite. Incorporating this information into the link is done by adding the cases to the prior parameter values. After updating, the link distribution is as follows:

$$\{P(R_1 | L_1), P(R_2 | L_1)\} \sim \text{Dirichlet}(417.5, 362.5).$$

Steere et. al.<sup>28</sup> reported that 97 early Lyme disease patients recalled a tickbite and 217 early Lyme patients didn't recall a tickbite. After updating the link for this information the link is as follows:

$$\{P(R_1 | L_1), P(R_2 | L_1)\} \sim \text{Dirichlet}(514.5, 579.5),$$

where the 97 patients in Steere's article are added to the prior parameter value of 417.5. No other journal articles were found where patients recalled being bitten by a tick. The initial estimate can now be calculated.

After all the information has been incorporated into the distribution the initial estimate can be computed. The Bayesian estimate is the mean of the posterior distribution. The mean of the dirichlet distribution (  $a_1, a_2, \dots, a_k$  ) is  $a_i / \sum_i a_i$  for  $i = 1, 2, \dots, k$ . For example, the initial estimate for the probability of recalling a tickbite given early Lyme is the first parameter divided by the sum of the parameters:

$$P(R_1 | L_1) = 514.5/1094$$

and the initial estimate for not recalling a tickbite is,

$$P(R_2 | L_1) = 579.5/1094.$$

Estimation of link 1 required the following alteration of the previously described method. The journal articles found reported data on the number of patients of each sex given Lyme disease and the number of patients with different months on onset given Lyme disease. This data yields estimates of the opposite conditional probabilities that are desired. Bayes theorem thus had to be applied. Bayes theorem can be written:

$$P(A|B) = P(B|A)*P(A) / P(B)$$

For example, the probability of early Lyme disease given a male patient is,

$$P(\text{early} | \text{male}) = \{P(\text{male} | \text{early})*P(\text{early})\} / P(\text{male}).$$

From table 1,  $P(\text{male} | \text{early}) = 0.4901$ ,  $P(\text{early Lyme disease}) = 0.0001$  and  $P(\text{male}) = 0.5$ . So, using Bayes theorem yields,

$$P(\text{early} | \text{male}) = (0.4901)*(0.0001)/(0.5) = 0.00009802.$$

It is assumed that the probability of being male given late Lyme disease is equal to the probability of being male given early Lyme disease, in order to do the calculations for late

Lyme. The probability of being male or female given no Lyme disease is calculated by noting,

$$\sum_i P(\text{male} | L_i)P(L_i) = P(\text{male}) = 1/2$$

where all quantities are known except  $P(\text{male} | L_0)$ .

### Section 3.2 Assumptions made:

Most journal articles used the patient inclusion criteria of present or previous ECM, a late sign or symptom of the disease and/or high titer levels from blood tests. To distinguish between early and late Lyme disease, this study considered patients with current ECM to have early Lyme disease and patients with characteristic late signs to have late Lyme patients. Late signs of Lyme disease include problems with the heart, joints, and nervous system, such as joint pain and swelling (arthritis), meningitis, Bell's palsy, cranial neuritis, and encephalitis.

The initial estimate for the probability of a patient recalling a tickbite given late Lyme disease was assumed to be the same as the estimate for early Lyme disease since late Lyme disease patients develop their symptoms soon after early infection. The tickbite that infected a patient with early Lyme disease will progress into late Lyme disease if not detected. Late Lyme disease doesn't result from an additional tickbite. Similarly, the initial estimate for the probability of recalling a tickbite given no Lyme disease and Fibromyalgia are the same because neither of these conditions increase the probability of recalling a tickbite.

The physician surveys were largely used to estimate probabilities given no Lyme disease. Any links not updated from the literature search were updated using the physician surveys. For example, the probability of a patient having a late sign of Lyme disease given he/she has early Lyme disease should be the same as not having Lyme disease at all.

The Bayesian belief network will use the initial estimates that were estimated in the literature search described in this chapter. All of the estimates are given in tables 1 through 18 in the appendix of this thesis. Each table includes the journal article that is being referenced, the number of patients in that article that have and don't have the respective symptom, and the initial estimate for the link in that table. Each table is headed with the condition of Lyme disease that it is estimating.

## **Chapter 4: Methods and results.**

In this chapter the Bayesian belief network is tested with data collected on case management forms from a number of South County physicians. For each patient, the expert system's diagnosis is compared to the serological diagnosis. An overall misclassification rate is calculated to determine how well the expert system classifies the patients. This chapter describes the case management forms, the classification process, and the results from testing the system.

### **4.1 Methods:**

Data on one hundred and twenty four patients was collected by three different medical offices in southern Rhode Island. The physician offices are Wood River Health Services in Hope Valley, Dr. Joseph England's office in Jamestown, and South County Physicians in Narragansett. Patients were included in the study if they had at least one symptom of Lyme disease. Patient symptoms were recorded in case management forms by the attending physician. The case management forms were collected by the Center for Vector-Borne Diseases at the University of Rhode Island. The case management form consisted of questions about demographic and historical variables, and clinical symptoms of Lyme disease. All of the physicians offices in this study were from an endemic area for Lyme disease. The study was intended to have physicians from both endemic and non-endemic regions of Rhode Island, however, presently no physicians from the non-endemic regions are enrolled.

The expert system yields a probability of Lyme disease given clinical symptoms of the disease. Even with strong evidence this probability is often small due to the small incidence rates. For example, the posterior probability of having early Lyme disease given a positive serum test is 0.0003 approximately three-fold higher than the incidence of 0.000127. Magid (1992) conducted a cost-benefit analysis of treating Lyme disease and determined that treatment is only warranted if the probability of Lyme disease is at least 0.01. Thus, we used a cut off level of 0.01 to classify a patient as having Lyme disease. If more than one state of Lyme disease or CFS was greater than 0.01 the patient was classified as the state with the largest likelihood.

The expert system developed used two computer packages called DXpress and Win-dx both by Knowledge Industries, Inc. DXpress built the tree of nodes and stored all the initial estimates for the links. Win-dx imported the causal tree and initial estimates from DXpress and calculated the posterior probabilities used for classification. These probabilities were rounded to six significant digits.

The serological results were provided by Imugen, a commercial testing laboratory. Imugen uses two tests: an antibody capture Enzyme-linked immunosorbent assay (ELISA) and a western blot analysis. Both estimate whether there are significant amounts of *Borrelia burgdorferi* antibody, an indicator of previous exposure or current infection with the Lyme disease bacterium. The ELISA tested for levels of three immunoglobulins: IgG, IgM, IgA. The western blot measured the number of bands produced against two antigens: G39/40 and FRG. Imugen reports a sensitivity of 0.91 and a specificity of 0.95 (Imugen pers. Comm.).

Imugen categorized patients into five states based on the serological results: negative, positive early, positive late, remote, and XR (cross reactivity). A classification of cross reactivity occurs when the antigen levels are present, but are not significantly high. This could indicate a very early response to Lyme disease or cross reactivity. Remote indicates a past infection rather than a current infection.

The study protocol was for each patient to have two serum tests. The first serum test would be administered immediately after the office visit and the second test would be given during a follow-up visit approximately two weeks later. Unfortunately, not all patients returned for the follow-up visit. All patients had at least one blood test.

Two different analyses of the expert system were done: first, the expert system's diagnosis based upon clinical symptoms was compared to Imugen's classification. Second, the expert system's diagnosis based upon clinical symptoms and the results of the first serum test were compared to Imugen's classification. The second comparison was conducted on only the patients who had two serum tests. An overall misclassification rate for the expert system's diagnosis was calculated along with specificity and sensitivity rates. This study classified patients using the results from both serum tests: at least one positive test resulted in a classification of Lyme disease.

#### **4.2 Results:**

Of 124 patients, 22 were classified as having Lyme disease and 102 were classified as not having Lyme disease by Imugen. No patients were classified as having CFS or



Fibromyalgia. The frequencies of the clinical symptoms for the 22 Lyme disease patients are as follows:

<u>Symptom</u>	<u>Early Lyme Number of Patients</u>	<u>Late Lyme Number of Patients</u>
<i>Myalgia</i>	4	5
<i>Arthralgia</i>	6	4
<i>Fever</i>	6	3
<i>Fatigue</i>	3	3
<i>Tender Points</i>	2	0
<i>Sleep Disorder</i>	1	1
<i>Headache</i>	3	5
<i>Stiff neck</i>	1	2
<i>Arthritis</i>	1	2

Of the 22 patients classified as having Lyme disease by the serum test, 14 had early Lyme and 8 had late Lyme. A patient was recorded as misclassified if the serum test classification and the classification by the expert system didn't match. 51 of the 124 patients were misclassified by the expert system using clinical variables alone and 18 of the 67 patients who had two serum tests were misclassified by the expert system when a serum test was included. The misclassification rates, specificities, and sensitivities are as follows:

	<u>Expert system without serum test</u>	<u>Expert system with serum test</u>
<i>Misclassification rate</i>	0.4113	0.2687
<i>Specificity</i>	0.6275	0.8163
<i>Sensitivity</i>	0.3636	0.5000

The classifications of all patients are listed in tables 19 through 21. The probabilities of each state in the disease node calculated by the expert system are displayed in table 22. Patients diagnosed by the expert system with CFS or Fibromyalgia were assumed to not have Lyme disease when calculating the misclassification rate.

The diagnostic accuracy of this first version of the expert system is a bit disappointing. In order to improve the system, various assumptions of the model were tested on the data. The next section describes in detail this analysis of the system.

#### **4.3 Analysis:**

Misclassified patients were separated into two groups in order to detect possible trends. Table 24 lists both patients that tested positive for Lyme but were classified by the expert system as not having Lyme and patients that tested negative for Lyme but were classified by the expert system as having Lyme. An examination of the first half of the table revealed that these patients had very few symptoms and that diagnosis on clinical symptoms alone would be very difficult. Five out of nine early Lyme patients had ECM, a fact which may indicate that the expert system should weight this symptom more heavily. An examination of the second half of the table revealed 30 patients that tested negative but were classified as having early Lyme disease by the expert system. These patients had a surprising number of symptoms of early Lyme disease. The following is a table of the symptoms experienced by these patients:

	<u>Number of patients</u>
<i>Recall tickbite</i>	17
<i>Myalgia</i>	20
<i>Arthralgia</i>	25
<i>Fever</i>	14
<i>Fatigue</i>	28
<i>Headache</i>	21
<i>Stiff neck</i>	22
<i>Tender points</i>	3
<i>Sleep disorder</i>	9

In order to assess whether this high number of symptoms was characteristic of Lyme patients, we compared these groups. The number of unmeasurable symptoms, i.e. myalgia, arthralgia, fatigue or malaise, headache, stiff neck, tender points, and sleep disorder, were recorded for each patient. The following table contains statistics on these symptoms.

	<u>Early Lyme</u>	<u>Late Lyme</u>	<u>No Lyme</u>
<i>Mean</i>	1.43	2.50	4.14
<i>median</i>	1.00	3.00	4.00
<i>upper quartile</i>	2.00	4.00	5.00
<i>lower quartile</i>	0.00	1.00	3.00
<i>N</i>	14.00	8.00	37.00

The group of 30 patients show a statistically significantly higher number of these symptoms than early Lyme patients. Persons with CFS have reported a high number of symptoms; perhaps these patient should be classified as CFS. In either case, a node with the number of symptoms may help the expert system distinguish this class of patients from patients with Lyme.

To test the initial estimates, a comparison of values from the literature search and the data on the 22 Lyme patients was done. Proportions of patients with symptoms in the network were calculated for early and late Lyme patients. These estimates were compared to the initial estimates from the literature search with a z-test for binomial proportions. Significant differences in the estimates of the proportions of early Lyme patients with headache, stiff neck, and fatigue or malaise, and in the estimates of the proportions of late Lyme patients with myalgia, arthralgia, and headache were found. The proportions for the 14 early Lyme patients and 8 late Lyme patients and the corresponding initial estimate from the literature search are:

	<i>Early Lyme Patients</i>		<i>Late Lyme Patients</i>	
	<u>proportion</u>	<u>initial estimate</u>	<u>proportion</u>	<u>initial estimate</u>
<i>Recall tickbite</i>	0.7143	0.4703	0.2500	0.4703
<i>Myalgia</i>	0.2857	0.4925	0.6250	0.2863*
<i>Arthralgia</i>	0.4286	0.3665	0.5000	0.0041*
<i>Fever</i>	0.4286	0.5999	0.3750	0.3542
<i>Fatigue</i>	0.2143	0.7764*	0.3750	0.4792
<i>Headache</i>	0.2143	0.5255*	0.6250	0.1734*
<i>Stiff neck</i>	0.0714	0.3552*	0.2500	0.2708
<i>Tender points</i>	0.1429	0.0950	0.0000	0.0950
<i>Sleep disorder</i>	0.0714	0.0934	0.1250	0.0934

\* denotes significant difference.

The highlighted symptoms have significant differences between the initial estimate and the proportions from the data. Biased estimates may lead to misdiagnoses.

One assumption of the Bayesian belief network is that symptoms are conditionally independent given the disease state. To test whether this is a valid assumption, Fisher's exact test for independence was run for each pair of clinical symptoms, conditioning on the disease state. The symptoms Bell's palsy, radiculoneuropathy, meningitis, encephalitis,

and nodal block weren't analyzed because of the small amount of patients having these symptoms.

Results from these tests indicated some dependencies among the clinical symptoms. Table 23 lists the significant pairs of symptoms for Lyme positive patients at a 0.05 significance level. There were 4 dependent pairs of symptoms for Lyme disease patients. These pairs of symptoms were myalgia and fever, arthralgia and fever, fatigue and fever, and headache and stiff neck. It is interesting to note that three of the pairs of symptoms consist of the three most frequent symptoms of the 22 Lyme patients. Myalgia and fever were both experienced by 9 of the 22 patients and arthralgia was experienced by 10 of the 22 patients. Dependent symptoms may cause misclassifications and tend to overdiagnose Lyme disease.

#### **4.4 Assessment of Physicians Diagnostic Ability:**

The case management forms queried each physician to estimate the likelihood that the given patient had Lyme disease, on a scale from 0-10. This assessment was based upon clinical symptoms alone. In this section we analyze the diagnostic abilities of our participating physicians by comparing the likelihoods given to patients classified as having Lyme by Imugen to those classified as not having Lyme. Statistics for the likelihoods are as follows:

	<b><u>Lyme patients</u></b>	<b><u>Non-Lyme patients</u></b>
<b>Mean</b>	6.50	4.67
<b>Variance</b>	7.74	5.03
<b>Median</b>	7.00	5.00
<b>Upper quartile</b>	9.50	6.50
<b>Lower quartile</b>	4.00	3.00

Figure 8 contains boxplots of the likelihood values for both Lyme and non-Lyme patients. These results show a higher mean likelihood for Lyme patients, but there is a significant amount of overlap in the two distributions. For accurate diagnosis, the distributions should have little overlap. In order to compare the physicians' diagnostic abilities to our expert system a misclassification rate, and specificity and sensitivity values were computed. We assumed a likelihood of zero to five classified a patient as not having Lyme; six to ten classified a patient as having Lyme. The misclassification, specificity, and sensitivity rates of these likelihoods are:

	<i><u>Physician Likelihoods</u></i>
<i>Misclassification rate</i>	0.3583
<i>Specificity</i>	0.6500
<i>Sensitivity</i>	0.6000

These rates are comparable with the expert system's diagnosis based on clinical symptoms. In fact, physicians at Wood River Health Services and Dr. England's office had exactly the same misclassification rate as the expert system; only physicians at South County did slightly better. The low sensitivity and specificity values highlight the difficulties in clinical diagnosis of Lyme disease.

A second comparison of the physicians' and expert system's clinical diagnoses was done with a regression analysis using a dependent variable of the physician's likelihood, an independent variable of the expert system's probability and no intercept, the R-squared value was 0.1497. This low correlation indicates that the expert system's diagnosis gives

little information about the likelihood. Figure 7 shows the data and the fitted regression line.

## **Chapter 5: Discussion and further research.**

The goal of this thesis is to develop an expert system to aid physicians in the diagnosis of Lyme disease. Toward that end, a literature search was done to calculate initial estimates for the expert system, and the system was tested on a sample of 124 patients. Misclassification, specificity, and sensitivity rates were calculated and proved to be surprisingly poor. An analysis of the system was conducted to find ways to improve its diagnostic ability.

We believe the system's poor performance was due to a number of factors. First, and foremost, clinical diagnosis of Lyme disease is very difficult, as evidenced by physicians' misclassification rates. Second, the system's diagnosis was compared to the serological diagnosis. Imugen's serum test is not a gold standard and the system's performance may have been better if patients were correctly classified. Third, a few of the initial estimates appeared to be incorrect. Finally, some clinical symptoms appeared to be conditionally dependent violating the assumptions of the Bayesian belief network.

We believe the system's accuracy can be substantially improved with a few simple solutions. The biased initial estimates should be reestimated. Some initial estimates for early Lyme patients appeared to be over estimated; whereas some values for late Lyme were underestimated. The reason for these biased estimates probably lies in the definitions of early versus late Lyme. Our definition did not always correspond to the definition in the articles. Some articles did not separate the stages; in these cases, we attempted to



divide the patients. To obtain better estimates, some of these articles should be thrown out of the study.

The Fisher's exact test used to test for independence revealed dependency between some symptoms. Four pairs of symptoms were significantly dependent, at the 0.05 significance level. Three of the pairs, myalgia and fever, arthralgia and fever, and fatigue and fever, interestingly contain the symptom fever. The fourth pair contains headache and stiff neck. Combining myalgia, arthralgia, fatigue, and fever into one node and headache and stiff neck in another node will alleviate the problems of conditional dependence.

The system could be further improved with the addition of new variables. Inclusion of exposure variables such as have you been bitten by a tick, was the tick infected, where do you reside, do you have pets, etc. should more accurately assess a patient's probability of having Lyme disease. The inclusion of a node that records the results of both the initial serum test and the follow-up serum test would be beneficial, since often early Lyme patients have not yet seroconverted at the time the first blood sample is taken. A node with the number of unmeasurable symptoms a patient experiences could help distinguish between Lyme disease and CFS.

Finally, the system's estimates can be refined by learning from the patient data. Diagnosis of Lyme disease has proven to be extremely challenging. The first version of our expert system performed slightly worse than physicians familiar with Lyme disease. We believe implementing the aforementioned changes will make the system surpass the physicians in diagnostic ability.

**Link1: P( month | early Lyme )**

	Bowen et.al. <sup>4</sup>	Steere et.al. <sup>35</sup>	Petersen, et.al. <sup>22</sup>	Hanrahan, et.al. <sup>15</sup>	CDC database	Row Total	INITIAL ESTIMATES
P(Jan.)	0	0	17	2	133	152.0833	0.0098
P(Feb.)	0	0	2	0	97	99.0833	0.0064
P(Mar.)	1	0	15	0	192	208.0833	0.0134
P(April)	4	1	30	2	439	476.0833	0.0307
P(May)	17	0	60	9	1309	1395.0833	0.0898
P(June)	31	4	157	35	4342	4569.0833	0.2942
P(July)	39	3	170	43	4771	5026.0833	0.3237
P(Aug.)	7	0	70	24	1705	1806.0833	0.1163
P(Sept.)	8	2	37	19	673	739.0833	0.0476
P(Oct.)	3	0	26	17	539	585.0833	0.0377
P(Nov.)	5	0	17	4	304	330.0833	0.0213
P(Dec.)	2	0	8	7	126	143.0833	0.0092
Column Total	117	10	609	162	14630	15529	

**Link1: P( month | late Lyme )**

	Steere et.al. <sup>35</sup>	CDC database	Row Total	INITIAL ESTIMATES
P(Jan.)	1	627	628.0833	0.0512
P(Feb.)	0	469	469.0833	0.0382
P(Mar.)	1	600	601.0833	0.0490
P(April)	1	773	774.0833	0.0631
P(May)	1	1128	1129.0833	0.0920
P(June)	1	1948	1949.0833	0.1588
P(July)	1	2159	2160.0833	0.1760
P(Aug.)	0	1391	1391.0833	0.1134
P(Sept.)	1	1012	1013.0833	0.0826
P(Oct.)	0	918	918.0833	0.0748
P(Nov.)	2	696	698.0833	0.0569
P(Dec.)	1	539	540.0833	0.0440
Column total	10	12260	12271	

**Link1:P( sex | early Lyme ) and P( sex | late Lyme )**

	Steere et.al. <sup>36</sup>	Bowen et.al. <sup>4</sup>	Steere et.al. <sup>38</sup>	Kaplan, et.al. <sup>18</sup>	Petersen, et.al. <sup>22</sup>	Dinerman and Steere <sup>8</sup>	CDC database	Row Total	INITIAL ESTIMATES
P(Male)	164	77	100	10	585	10	9319	10265.5	0.4901
P(Female)	150	40	80	10	564	5	9831	10680.5	0.5099
Column Total	314	117	180	20	1149	15	19150	20946	

**Link1: P( sex | CFS or Fibromyalgia )**

	Goldenberg <sup>11</sup>	Sigal and Patella <sup>31</sup>	Steere et.al. <sup>38</sup>	Kaplan, et.al. <sup>18</sup>	Whelton et.al. <sup>41</sup>	Buchwald and Garrity <sup>5</sup>	Hsu et.al. <sup>16</sup>	Row Total	INITIAL ESTIMATES
P(Male)	40	8	88	2	3	15	13	169.5	0.2391
P(Female)	205	22	183	9	11	45	64	539.5	0.7609
Column Total	245	30	271	11	14	60	77	709	

Table 1: Link 1 Initial estimates of the probability of sex and month of onset given Lyme disease.

**Link1: P( Lyme disease | Month of Onset and Sex )**

	<b>Jan.</b>	<b>Feb.</b>	<b>March</b>	<b>April</b>	<b>May</b>	<b>June</b>
<b>No Lyme</b>	0.981862	0.981897	0.981863	0.981809	0.981672	0.981274
<b>Early Lyme</b>	0.000012	0.000008	0.000016	0.000037	0.000108	0.000353
<b>Late Lyme</b>	0.000123	0.000092	0.000118	0.000152	0.000221	0.000381
<b>CFS</b>	0.018003	0.018004	0.018003	0.018002	0.017999	0.017992

	<b>July</b>	<b>Aug.</b>	<b>Sept.</b>	<b>Oct.</b>	<b>Nov.</b>	<b>Dec.</b>
<b>No Lyme</b>	0.981198	0.981590	0.981744	0.981774	0.981835	0.981880
<b>Early Lyme</b>	0.000388	0.000139	0.000057	0.000045	0.000026	0.000011
<b>Late Lyme</b>	0.000422	0.000272	0.000198	0.000179	0.000137	0.000106
<b>CFS</b>	0.017991	0.017998	0.018001	0.018001	0.018002	0.018003

Table 1(cont.): Link 1 Initial estimates of the probability of Lyme disease given sex and month of onset.

**Link2: P( recall tickbite / no Lyme )**

	Petersen, et.al. <sup>22</sup>	Row Total	INITIAL ESTIMATES
<b>P(recall tickbite)</b>	360	360.5	0.3102
<b>P(don't recall tickbite)</b>	801	801.5	0.6898
Column Total	1161	1162	

**Link2: P( recall tickbite / early Lyme )**

	Steere et. al. <sup>35</sup>	Petersen, et.al. <sup>22</sup>	Row Total	INITIAL ESTIMATES
<b>P(recall tickbite)</b>	97	417	514.5	0.4703
<b>P(don't recall tickbite)</b>	217	362	579.5	0.5297
Column Total	314	779	1094	

**Link2: P( recall tickbite / late Lyme )**

	Steere et. al. <sup>35</sup>	Petersen, et.al. <sup>22</sup>	Row Total	INITIAL ESTIMATES
<b>P(recall tickbite)</b>	97	417	514.5	0.4703
<b>P(don't recall tickbite)</b>	217	362	579.5	0.5297
Column Total	314	779	1094	

**Link2: P( recall tickbite / CFS or Fibromyalgia )**

	Petersen, et.al. <sup>22</sup>	Row Total	INITIAL ESTIMATES
<b>P(recall tickbite)</b>	360	360.5	0.3102
<b>P(don't recall tickbite)</b>	801	801.5	0.6898
Column Total	1161	1162	

Table 2: Link 2 Initial estimates of the probability of recalling a tickbite given Lyme disease.

**Link3: P( myalgia | no Lyme )**

	Survey data	Waylonis and Heck <sup>40</sup>	Row Total	INITIAL ESTIMATES
<b>P(Myalgia)</b>	72	30	102.5	0.1309
<b>P(No Myalgia)</b>	541	139	680.5	0.8691
Column total	613	169	783	

**Link3: P( myalgia | early Lyme )**

	Steere et.al. <sup>36</sup>	Steere et.al. <sup>34</sup>	Hanrahan, et.al. <sup>15</sup>	Petersen et.al. <sup>22</sup>	Dineman and Steere <sup>8</sup>	Dekonenko <sup>7</sup>	Row Total	INITIAL ESTIMATES
<b>P(Myalgia)</b>	135	3	143	181	6	24	492.5	0.4925
<b>P(No Myalgia)</b>	179	9	55	189	9	66	507.5	0.5075
Column total	314	12	198	370	15	90	1000	

**Link3: P( myalgia | late Lyme )**

	Wahlberg et. al. <sup>39</sup>	Steere et.al. <sup>34</sup>	Row Total	INITIAL ESTIMATES
<b>P(Myalgia)</b>	32	3	35.5	0.2863
<b>P(No Myalgia)</b>	68	20	88.5	0.7137
Column total	100	23	124	

**Link3: P( myalgia | CFS or Fibromyalgia )**

	Buchwald and Garrity <sup>5</sup>	Row Total	INITIAL ESTIMATES
<b>P(Myalgia)</b>	52	52.5	0.8607
<b>P(No Myalgia)</b>	8	8.5	0.1393
Column total	60	61	

Table 3: Link 3 Initial estimates of the probability of myalgia given Lyme disease.

**Link4: P( arthralgia | no Lyme ) and P( arthralgia | late Lyme )**

	Survey data	Row Total	INITIAL ESTIMATES
<b>P(Arthralgia)</b>	2	2.5	0.0041
<b>P(No Arthralgia)</b>	611	611.5	0.9959
Column total	613	614	

**Link4: P( arthralgia | early Lyme )**

	Steere et.al. <sup>36</sup>	Bowen et.al. <sup>4</sup>	Petersen, et.al. <sup>22</sup>	Dinerman and Steere <sup>8</sup>	Dekonenko <sup>7</sup>	Row Total	INITIAL ESTIMATES
<b>P(Arthralgia)</b>	150	24	126	7	22	329.5	0.3665
<b>P(No Arthralgia)</b>	164	84	245	8	68	569.5	0.6335
Column total	314	108	371	15	90	899	

**Link4: P( arthralgia | CFS or Fibromyalgia )**

	Wahlberg <sup>39</sup>	Buchwald and Garrity <sup>5</sup>	Row Total	INITIAL ESTIMATES
<b>P(Arthralgia)</b>	11	49	60.5	0.3758
<b>P(No Arthralgia)</b>	89	11	100.5	0.6242
Column total	100	60	161	

Table 4: Link 4 Initial estimates of the probability of arthralgia given Lyme disease.

**Link5: P( fever | no Lyme )**

	Survey data	Row Total	INITIAL ESTIMATES
<b>P(Fever)</b>	54	54.5	0.0888
<b>P(No Fever)</b>	559	559.5	0.9112
Column total	613	614	

**Link5: P( fever | early Lyme )**

	Steere et.al. <sup>36</sup>	Steere et.al. <sup>34</sup>	Hanrahan, et.al. <sup>15</sup>	Petersen et.al. <sup>22</sup>	Dinerman and Steere <sup>8</sup>	Row Total	INITIAL ESTIMATES
<b>P(Fever)</b>	185	9	129	214	12	549.5	0.5999
<b>P(No Fever)</b>	129	3	69	162	3	366.5	0.4001
Column total	314	12	198	376	15	916	

**Link5: P( fever | late Lyme )**

	Steere et.al. <sup>34</sup>	Row Total	INITIAL ESTIMATES
<b>P(Fever)</b>	8	8.5	0.3542
<b>P(No Fever)</b>	15	15.5	0.6458
Column total	23	24	

**Link5: P( fever | CFS or Fibromyalgia )**

	Buchwald and Garrity <sup>5</sup>	Row Total	INITIAL ESTIMATES
<b>P(Fever)</b>	20	20.5	0.3361
<b>P(No Fever)</b>	40	40.5	0.6639
Column total	60	61	

Table 5: Link 5 Initial estimates of the probability of fever given Lyme disease.

**Link6: P( fatigue or malaise / no Lyme )**

	Survey data	Waylonis, Heck et.al. <sup>40</sup>	Row Total	INITIAL ESTIMATES
P(Fat.or Mal.)	45	30	75.5	0.0964
P(No Fat.or Mal.)	568	139	707.5	0.9036
Column total		169	783	

**Link6: P( fatigue or malaise / early Lyme )**

	Hanrahan, et.al. <sup>15</sup>	Steere et.al. <sup>36</sup>	Steere et.al. <sup>28</sup>	Dekonenko <sup>7</sup>	Row Total	INITIAL ESTIMATES
P(Fat.or Mal.)	150	251	9	67	477.5	0.7764
P(No Fat.or Mal.)	48	63	3	23	137.5	0.2236
Column total	198	314	12	90	615	

**Link6: P( fatigue or malaise / late Lyme )**

	Steere et.al. <sup>34</sup>	Row Total	INITIAL ESTIMATES
P(Fat.or Mal.)	11	11.5	0.4792
P(No Fat.or Mal.)	12	12.5	0.5208
Column total	23	24	

**Link6: P( fatigue or malaise ICFS or Fibromyalgia )**

	Goldenberg <sup>11</sup>	Sigal and Patella <sup>31</sup>	Goldenberg et.al. <sup>13</sup>	Buchwald and Garrity <sup>5</sup>	Hsu et.al. <sup>16</sup>	Row Total	INITIAL ESTIMATES
P(Fat.or Mal.)	224	26	45	58	69	422.5	0.9185
P(No Fat.or Mal.)	21	4	2	2	8	37.5	0.0815
Column total	245	30	47	60	77	460	

Table 6: Link 6 Initial estimates of the probability of malaise or fatigue given Lyme disease.



**Link7: P(tender pts. | no Lyme ), P(tender pts. | early Lyme ), and P(tender pts. | late Lyme )**

	Survey data	Row Total	INITIAL ESTIMATES
<b>P(Tender Points)</b>	58	58.5	0.095
<b>P(No Tender Points)</b>	555	555.5	0.905
Column total		614	

**Link7: P(tender pts. | CFS or fibromyalgia )**

	Sigal and Patella <sup>31</sup>	Hsu et.al. <sup>16</sup>	Row Total	INITIAL ESTIMATES
<b>P(Tender Points)</b>	30	107	137.5	0.996
<b>P(No Tender Points)</b>	0	0	0.5	0.004
Column total	30	107	138	

Table 7: Link 7 Initial estimates of the probability of tender points given Lyme disease.

**Link8: P( stiff neck | no Lyme ) and P( stiff neck | CFS or fibromyalgia )**

	Survey data	Row Total	INITIAL ESTIMATES
<b>P(Stiff Neck)</b>	9	9.5	0.0155
<b>P(No Stiff Neck)</b>	604	604.5	0.9845
Column total	613	614	

**Link8: P( stiff neck | early Lyme )**

	Steere et.al. <sup>36</sup>	Steere et.al. <sup>34</sup>	Hanrahan, et.al. <sup>15</sup>	Petersen et.al. <sup>22</sup>	Dinerman and Steere <sup>8</sup>	Row Total	INITIAL ESTIMATES
<b>P(Stiff Neck)</b>	151	6	82	75	8	322.5	0.3552
<b>P(No Stiff Neck)</b>	163	6	109	300	7	585.5	0.6448
Column total	314	12	191	375	15	908	

**Link8: P( stiff neck | late Lyme )**

	Steere et.al. <sup>34</sup>	Row Total	INITIAL ESTIMATES
<b>P(Stiff Neck)</b>	6	6.5	0.2708
<b>P(No Stiff Neck)</b>	17	17.5	0.7292
Column total	23	24	

Table 8: Link 8 Initial estimates of the probability having a stiff neck given Lyme disease.

**Link9: P( headache | no Lyme )**

	Survey data	Row Total	INITIAL ESTIMATES
P(Headache)	72	72.5	0.1181
P(No Headache)	541	541.5	0.8819
Column total	613	614	

**Link9: P( headache | early Lyme )**

	Steere et.al. <sup>35</sup>	Steere et.al. <sup>34</sup>	Hanrahan, et.al. <sup>15</sup>	Petersen, et.al. <sup>22</sup>	Dinerman and Steere <sup>8</sup>	Dekonenko <sup>7</sup>	Row Total	INITIAL ESTIMATES
P(Headache)	200	6	120	164	14	52	556.5	0.5255
P(No Headache)	114	106	35	209	1	38	503.5	0.4754
Column total	314	112	155	373	15	90	1059	

**Link9: P( headache | late Lyme )**

	Wahlberg et. al. <sup>30</sup>	Steere et.al. <sup>34</sup>	Row Total	INITIAL ESTIMATES
P(Headache)	13	8	21.5	0.1734
P(No Headache)	87	15	102.5	0.8266
Column total	100	23	124	

**Link9: P( headache | CFS or Fibromyalgia )**

	Goldenberg <sup>11</sup>	Wheaton et.al. <sup>41</sup>	Sigal and Patella <sup>31</sup>	Buchwald and Garrity <sup>2</sup>	Hsu et.al. <sup>18</sup>	Row Total	INITIAL ESTIMATES
P(Headache)	131	36	24	45	42	278.5	0.6094
P(No Headache)	114	8	6	15	35	178.5	0.3906
Column total	245	44	30	319	77	457	

Table 9: Link 9 Initial estimates of the probability of headache given Lyme disease.

**Link10: P( sleep disorder | no Lyme )**

	Survey data	Waylonis & Heck <sup>40</sup>	Row Total	INITIAL ESTIMATES
P(Sleep Disorder)	5	35	40.5	0.0517
P(No Sleep Disorder)	608	134	742.5	0.9483
Column total	613	169	783	

**Link10: P( sleep disorder | early Lyme ) and P( sleep disorder | late Lyme )**

	Dekonenko <sup>7</sup>	Row Total	INITIAL ESTIMATES
P(Sleep Disorder)	8	8.5	0.0934
P(No Sleep Disorder)	82	82.5	0.9066
Column total	90	91	

**Link10: P( sleep disorder | CFS or Fibromyalgia)**

	Golden-berg <sup>11</sup>	Sigal & Patella <sup>31</sup>	Goldenberg et.al. <sup>13</sup>	Buchwald and Garrity <sup>5</sup>	Hsu et.al. <sup>16</sup>	Dinerman & Steere <sup>8</sup>	Row Total	INITIAL ESTIMATES
P(Sleep Disorder)	173	30	43	39	77	14	376.5	0.7926
P(No Sleep Disorder)	72	0	4	21	0	1	98.5	0.2074
Column total	245	30	47	60	77	15	475	

Table 10: Link 10 Initial estimates of the probability of sleep disorder given Lyme disease.

**Link11: P(Meningitis | no Lyme ), P(Meningitis | early Lyme ), and P(Meningitis | CFS or Fibromyalgia )**

	Survey data	Row Total	INITIAL ESTIMATES
<b>P(Meningitis)</b>	0	0.5	0.0008
<b>P(No Meningitis)</b>	613	613.5	0.9992
Column total	613	614	

**Link11: P(Meningitis | late Lyme )**

	Bowen et.al. <sup>4</sup>	Petersen et.al. <sup>22</sup>	Wahlberg et. al. <sup>39</sup>	CDC database	Dinerman and Steere <sup>8</sup>	Row Total	INITIAL ESTIMATES
<b>P(Meningitis)</b>	12	14	10	330	3	369.5	0.0355
<b>P(No Meningitis)</b>	37	344	90	9567	12	10050.5	0.9645
Column total	49	358	100	9897	15	10420	

Table 11: Link 11 Initial estimates of the probability of meningitis given Lyme disease.

**Link12: P( radiculoneuropathy | no Lyme ) and P( radiculoneuropathy | CFS or fibromyalgia )**

	Survey data	Row Total	INITIAL ESTIMATES
P(radiculoneuropathy)	0	0.5	0.0008
P(No radiculoneuropathy)	613	613.5	0.9992
Column total	613	614	

**Link12: P( radiculoneuropathy | early Lyme )**

	Bowen et.al. <sup>4</sup>	Petersen et.al. <sup>22</sup>	Row Total	INITIAL ESTIMATES
P(radiculoneuropathy)	0	0	0.5	0.0004
P(No radiculoneuropathy)	117	1111	1228.5	0.9996
Column total	117	1111	1229	

**Link12: P( radiculoneuropathy | late Lyme )**

	Wahlberg et. al. <sup>39</sup>	Dinerman and Steere <sup>8</sup>	Row Total	INITIAL ESTIMATES
P(radiculoneuropathy)	0	0	0.5	0.0043
P(No radiculoneuropathy)	100	15	115.5	0.9957
Column total	100	15	116	

Table 12: Link 12 Initial estimates of the probability of radiculoneuropathy given Lyme disease.

**Link13: P(Encephalitis | no Lyme ), P(Encephalitis | early Lyme ), and P(Encephalitis | CFS or Fibromyalgia )**

	Survey data	Row Total	INITIAL ESTIMATES
<b>P(Encephalitis)</b>	0	0.5	0.0008
<b>P(No Encephalitis)</b>	613	613.5	0.9992
Column total	613	614	

**Link13: P(Encephalitis | late Lyme )**

	Bowen et.al. <sup>4</sup>	Petersen et.al. <sup>22</sup>	Wahlberg et. al. <sup>39</sup>	CDC database	Dinerman and Steere <sup>8</sup>	Row Total	INITIAL ESTIMATES
<b>P(Encephalitis)</b>	1	14	0	232	0	247.5	0.0296
<b>P(No Encephalitis)</b>	48	344	100	7601	15	8108.5	0.9704
Column total	49	358	100	7833	15	8356	

Table 13: Link 13 Initial estimates of the probability of encephalitis given Lyme disease.

**Link14: P(Bell's palsy | no Lyme ), P(Bell's palsy | early Lyme ), and P(Bell's palsy | CFS or Fibromyalgia )**

	Survey data	Row Total	INITIAL ESTIMATES
<b>P(Bell's Palsy)</b>	0	0.5	0.0008
<b>P(No Bell's Palsy)</b>	613	613.5	0.9992
Column total	613	614	

**Link14: P(Bell's palsy | late Lyme )**

	Bowen et.al. <sup>4</sup>	Petersen et.al. <sup>22</sup>	Wahlberg et. al. <sup>39</sup>	Hanrahan, et.al. <sup>15</sup>	CDC database	Dinerman and Steere <sup>8</sup>	Row Total	INITIAL ESTIMATES
<b>P(Bell's Palsy)</b>	9	29	12	34	1748	4	1836.5	0.1528
<b>P(No Bell's Palsy)</b>	40	329	88	157	9559	11	10184.5	0.8472
Column total	49	358	100	191	11307	15	12021	

Table 14: Link 14 Initial estimates of the probability of Bell's palsy given Lyme disease.



**Link15: P(Nodal block | no Lyme ), P(Nodal block | early Lyme ), and P(Nodal block | CFS or Fibromyalgia )**

	Survey data	Row Total	INITIAL ESTIMATES
<b>P(Nodal Block)</b>	0	0.5	0.0008
<b>P(No Nodal Block)</b>	613	613.5	0.9992
Column total	613	614	

**Link15: P(Nodal block | late Lyme )**

	CDC database	Row Total	INITIAL ESTIMATES
<b>P(Nodal Block)</b>	224	224.5	0.0265
<b>P(No Nodal Block)</b>	8237	8237.5	0.9735
Column total	8461	8462	

Table 15: Link 15 Initial estimates of the probability of nodal block given Lyme disease.

**Link16: P(arthritis | no Lyme and child ) and P(arthritis | early Lyme and child )**

	Survey data	Row Total	INITIAL ESTIMATES
P(Arthritis)	0	0.5	0.0030
P(No Arthritis)	167	167.5	0.9970
Column total	167	168	

**Link16: P(arthritis | late Lyme and child )**

	Petersen et.al. <sup>22</sup>	CDC database	Row Total	INITIAL ESTIMATES
P(Arthritis)	113	2497	2610.5	0.4325
P(No Arthritis)	30	3395	3425.5	0.5675
Column total	143	5892	6036	

**Link16: P(arthritis | CFS or Fibromyalgia and child )**

	Sigal and Patella <sup>31</sup>	Row Total	INITIAL ESTIMATES
P(Arthritis)	0	0.5	0.0161
P(No Arthritis)	30	30.5	0.9839
Column total	30	31	

**Link16: P(arthritis | no Lyme and adult ), P(arthritis | early Lyme and adult ), and P(arthritis | CFS or Fibromyalgia and adult )**

	Survey data	Row Total	INITIAL ESTIMATES
P(Arthritis)	0	0.5	0.0011
P(No Arthritis)	446	446.5	0.9989
Column total	446	447	

**Link16: P(arthritis | late Lyme and adult )**

	Petersen et.al. <sup>22</sup>	Wahlberg et. al. <sup>39</sup>	Dinerman and Steere <sup>8</sup>	CDC database	Row Total	INITIAL ESTIMATES
P(Arthritis)	129	27	6	9213	9376	0.5363
P(No Arthritis)	68	73	9	7955	8106	0.4637
Column total	197	100	15	17168	17481	

Table 16: Link 16 Initial estimates of the probability of oligoarticular arthritis given Lyme disease and age.

<b>Link 17 Initial Estimates</b>				
<b>Link17: P(ECM   Lyme disease )</b>				
	No Lyme	Early Lyme	Late Lyme	CFS
<b>P(ECM)</b>	0.05	0.79	0.07	0.05
<b>P(No ECM)</b>	0.94	0.20	0.20	0.94
<b>P(Previous ECM)</b>	0.01	0.01	0.73	0.01

Table 17: Link 17 Initial estimates of the probability of ECM given Lyme disease.

**Link18: P(serum test | no Lyme ) and P(serum test | CFS or Fibromyalgia )**

	Imugen	Row Total	INITIAL ESTIMATES
P(Remote)	0	0.2	0.0333
P(Positive,early)	0	0.2	0.0333
P(Positive, late)	0	0.2	0.0333
P(Negative)	4	4.2	0.7000
P(XR)	1	1.2	0.2000
Column Total	5	6	

**Link18: P(serum test | early Lyme )**

	Imugen	Row Total	INITIAL ESTIMATES
P(Remote)	0	0.2	0.0087
P(Positive,early)	12	12.2	0.5304
P(Positive, late)	0	0.2	0.0087
P(Negative)	2	2.2	0.0957
P(XR)	8	8.2	0.3565
Column Total	22	23	

**Link18: P(serum test | late Lyme )**

	Imugen	Row Total	INITIAL ESTIMATES
P(Remote)	0	0.2	0.0133
P(Positive,early)	1	1.2	0.0800
P(Positive, late)	13	13.2	0.8800
P(Negative)	0	0.2	0.0133
P(XR)	0	0.2	0.0133
Column Total	14	15	

Table 18: Link 18 Initial estimates of the probability of a first serum test given Lyme disease

Wood River Health Services

Patient #	Sex	Onset	Recall bits	Myalgia	Arthralgia	Fever	Fat.or Mal.	Headache	Stiff neck	Tender Pts.	Sleep dis.	Bells palsy	radiculo.	Meningitis	Enceph.	Nodal	Arthritis	ECM	2 serum	1 serum	Age	Treated	Likelihood	TRUE Class#	Expert w/o serum	Expert w/ serum	
501	m	9	n	n	y	n	n	n	n	n	n	n	n	n	n	n	y	n	neg,neg		a	n	5	no	no	no	
502	f	8	n	y	y	n	y	y	n	n	n	n	n	n	n	n	n	n		neg	c	n	6	no	early	cts	
503	m	10	y	n	n	y	y	y	n	n	n	n	n	n	n	n	n	n		neg	a	n	3	no	no	no	
504	f	5	y	y	y	n	y	y	y	n	n	n	y	n	n	n	n	n	neg,neg		a	n	4	no	early	cts	
505	m	4	y	n	y	n	y	n	n	y	n	n	n	n	n	n	n	n		neg	a	n	3	no	cts	cts	
506	m	10	y	y	y	n	n	y	n	n	n	n	n	n	n	n	n	n	neg,neg		a	n	2	no	no	no	
507	m	10	y	n	y	n	y	y	n	n	n	n	y	n	n	n	n	n	XR,neg		a	y	7	no	no	no	
508	m	10	n	y	y	y	y	y	n	y	n	n	n	n	n	n	n	n	neg,neg		a	n	5	no	cts	cts	
509	f	10	n	y	y	n	y	y	y	y	n	n	n	n	n	n	n	n	neg,neg		c	n	2	no	cts	cts	
510	f	12	n	y	y	y	y	y	n	n	n	y	n	n	n	n	n	n		neg	a	y	8	no	cts	cts	
511	m	10	y	y	y	n	y	y	y	y	n	n	n	n	n	n	n	n		neg	a	n	4	no	cts	cts	
512	m	10	y	y	y	y	y	y	n	n	y	n	n	n	n	n	n	n		neg	a	y	8	no	cts	cts	
513	m	11	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	neg,neg		a	y	6	no	no	no	
514	f	11	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	PE,PE		a	n	4	early	no	no	
515	f	11	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	y	neg,PE		a	y	10	early	no	no	
516	f	11	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	neg,neg		a	y	5	no	no	no	
517	f	11	y	n	y	y	y	y	n	n	n	n	n	n	n	n	n	y	neg,neg		c	y	7	no	early	early	
518	m	10	y	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	PE,PE		c	n	5	early	no	early	
519	m	11	n	n	y	n	n	n	n	n	n	n	n	n	n	n	y	n	PL,PL		a	y	7	late	no	late	
520	m	11	n	y	y	n	y	y	y	n	n	n	n	n	n	n	n	n	neg,neg		a	n	4	no	early	early	
521	f	12	y	y	y	n	y	y	y	n	n	n	n	n	n	n	n	n	neg,neg		a	y	5	no	early	early	
522	f	11	y	y	y	y	y	y	n	y	n	n	n	n	n	n	n	n	prev	neg,neg		a	y	6	no	early	cts
523	f	12	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	PE,PE		c	n	2	early	no	no	
524	f	12	y	y	n	y	y	y	y	n	n	n	n	n	n	n	n	n		neg	c	y	8	no	early	no	
525	f	12	y	y	y	n	y	y	y	y	n	n	n	n	n	n	n	n		neg	a	n	7	no	cts	cts	
526	m	3	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	XR,PE		a	n	2	early	no	no	
527	m	3	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	neg,neg		a	n	2	no	no	no	
528	m	3	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y	neg,neg		a	n	4	no	no	no	
529	m	4	n	y	y	n	y	n	n	n	n	n	n	n	n	n	n	n	neg,neg		a	y	5	no	no	no	
530	m	4	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	neg,neg		a	n	3	no	no	no	
531	f	2	n	n	y	y	y	y	n	n	n	n	n	n	n	n	y	n	neg,neg		a	y	9	no	late	no	
532	f	4	n	y	y	n	y	y	y	n	y	n	n	n	n	n	n	n	XR,XR		a	n	3	no	early	early	
533	m	4	n	y	y	y	y	y	n	n	n	n	n	n	n	n	n	n	PL,PL		a	y	7	late	early	early	
534	m	4	n	y	y	n	y	n	y	n	n	n	n	n	n	n	n	n	neg,neg		a	n	3	no	early	early	
535	m	5	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	neg,neg		a	n	4	no	no	no	
536	m	12	y	n	y	n	y	n	y	n	n	n	n	n	n	n	n	n	neg,neg		a	n	3	no	no	no	
537	m	5	y	n	y	n	y	n	n	n	n	n	n	n	n	n	n	n	neg,neg		a	n	2	no	no	no	
538	m	4	y	y	y	y	n	y	n	n	n	n	n	n	n	n	n	y		PE	a	y	10	early	early	early	
539	m	5	n	y	y	y	y	y	n	y	n	n	n	n	n	n	n	n		neg	c	y	7	no	cts	cts	
540	f	5	n	n	n	n	n	n	n	n	y	n	n	n	n	n	y	n	neg,neg		a	n	3	no	no	no	
541	m	5	y	n	n	n	n	n	n	y	n	n	n	n	n	n	n	y	PE,PE		a	y	8	early	no	no	
542	m	5	n	n	y	n	n	y	n	n	n	n	n	n	n	n	y	n	neg,neg		a	n	5	no	late	cts	
543	f	5	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	neg,neg		a	n	2	no	no	no	
544	m	4	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n		neg	a	n	8	no	no	no	
545	f	5	n	n	n	n	n	n	n	n	y	n	n	n	n	n	n	y	neg,neg		a	y	8	no	no	no	
546	f	5	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	neg,PE		a	y	9	early	no	no	
547	f	6	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n		neg	c	n	2	no	no	no	
548	m	5	y	y	y	n	y	n	y	n	y	n	n	n	n	n	n	n	neg,neg		a	y	5	no	early	early	
549	m	6	n	y	y	y	y	n	y	n	y	n	n	n	n	n	n	n	prev	neg	a	n	4	no	early	early	
550	f	7	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y		PE	c	y		early	no	no	
551	m	6	y	y	n	y	y	n	y	n	n	n	n	n	n	n	n	n		neg	a	y	8	no	early	early	
552	f	6	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	neg,neg		a	y	8	no	no	no	
553	m	6	y	y	n	n	y	y	n	n	n	n	n	n	n	n	n	n		neg	a	y	6	no	no	no	
554	f	6	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	y	rem,rem		a	y	8	no	no	no	

Table 19: Patient data from Wood River Health Services

Patient #	Sex	Onset	Recall site	Myalgia	Arthralgia	Fever	Fat. or Mal.	Headache	Stiff neck	Tender Pts.	Strep obs.	Bedle palsy	radiculo.	Meningitis	Enceph.	Nodal	Arthritis	ECM	2 serum	1 serum	Age	Treated	Likelihood	TRUE Classif.	Expert w/ serum	Expert w/ serum
601	f	6	Y	Y	n	n	n	Y	Y	n	n	Y	n	n	n	n	n	n	PL, PL		a	Y	7	late	no	late
602	m	6	n	Y	n	n	Y	Y	n	n	n	n	n	n	n	n	n	Y	neg/neg		a	Y	7	no	early	early
603	m	7	n	Y	Y	Y	Y	n	Y	n	n	n	n	n	n	n	n	n	rem/rem		n	n	2	no	no	no
604	m	7	n	Y	Y	Y	Y	n	n	n	Y	n	n	n	n	n	Y	n	PE/PE	PL	c	n	4	late	early	late
605	f	8	n	n	n	n	n	Y	n	n	n	Y	n	n	n	n	n	n	PE/PE		a	Y	4	late	no	no
606	f	7	n	Y	n	n	Y	Y	n	n	n	n	n	n	n	n	n	n	PE/PE	PL	a	Y	4	late	no	no
607	f	8	Y	Y	n	n	Y	Y	n	n	n	n	n	n	n	n	n	Y	PE/PE		c	Y	10	early	early	early
608	f	8	n	n	n	Y	n	Y	n	n	n	n	n	n	n	n	n	n	neg	neg	c	Y	7	no	no	no
609	f	8	n	n	n	n	Y	Y	n	n	Y	n	n	n	n	n	n	n	neg/neg		a	Y	6	no	no	no
610	f	8	n	n	n	n	Y	Y	n	n	n	n	n	n	n	n	n	n	neg/neg		a	Y	1	no	no	no
611	f	7	Y	Y	n	n	Y	Y	n	n	n	n	n	n	n	n	n	Y	PL, PL		a	Y	6	late	early	early
612	m	9	n	n	n	Y	Y	Y	n	n	n	n	n	n	n	n	n	n	neg/neg		a	n	1	no	no	no
613	m	9	n	n	n	Y	Y	n	n	n	n	n	n	n	n	n	n	n	neg/neg		a	n	3	no	no	no
614	m	9	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	neg		c	Y	3	no	no	no
615	m	9	n	n	n	n	Y	Y	Y	n	n	n	n	n	n	n	n	n	neg		a	Y	3	no	no	no
616	m	10	n	n	n	n	Y	Y	n	n	n	n	n	n	n	n	n	n	neg/neg		a	n	4	no	early	no
617	m	10	Y	Y	n	n	n	Y	n	n	n	n	n	n	n	n	n	n	neg/neg		a	n	1	no	no	no
618	n	9	n	n	n	n	Y	n	Y	n	n	n	n	n	n	n	n	n	neg		a	n	2	no	no	no
619	n	10	Y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	neg		a	Y	10	early	no	no
620	n	11	n	n	n	n	Y	n	n	n	n	n	n	n	n	n	n	Y	XRF/PE		a	Y	2	no	no	no
622	m	9	n	Y	Y	Y	Y	n	n	n	n	n	n	n	n	n	n	n	PE/rem		a	Y	3	early	early	early
623	f	5	n	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	n	neg		a	n	4	no	cls	cls
624	f	5	n	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	n	neg/neg		a	n	2	no	cls	cls
625	f	6	Y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	neg		a	n	1	no	no	no
626	f	6	Y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	Y	neg/neg		a	Y	8	no	no	no
627	f	6	n	n	n	n	n	Y	n	n	n	n	n	n	n	n	n	Y	PL, PL		c	Y	7	late	no	no
628	m	6	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	prev		a	Y	2	no	no	no
629	f	12	n	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	n	neg		a	n	8	no	cls	cls

Table 20 Patient data from Dr. Joseph England's office

South County Physicians

Patient #	Sex	Onset	Recall bite	Myalgia	Arthralgia	Fever	Fat. or Mal.	Headache	Stiff neck	Tender Pts.	Sleep dis.	Bells palsy	radiculo.	Meningitis	Enceph.	Nodal	Arthritis	ECM	2 serum	1 serum	Age	Treated	Likelihood	TRUE Classif.	Expert w/o serum	Expert w/ serum	
1101	m	9	n	y	n	y	y	n	n	n	n	n	n	n	n	n	n	n		neg	a	y	4	no	no	no	
1102	m	9	n	n	n	n	y	n	n	n	y	n	n	n	n	n	n	n		neg	a			no	no	no	
1103	f	8	n	n	n	y	y	y	y	y	n	n	n	n	n	n	n	n	neg, neg		a	y	3	no	early	cts	
1104	m	9	y	n	y	n	y	y	n	n	n	n	n	n	n	n	n	n	neg, neg		a	n	5	no	no	no	
1105	m	10	n	n	n	n	y	y	n	n	n	n	n	n	n	n	n	n	neg, neg		a	y	2	no	no	no	
1106	m	10	n	n	n	n	y	n	n	n	n	n	n	n	n	n	n	n	neg, neg		a	n	2	no	no	no	
1107	f	9	n	n	y	n	n	y	n	n	n	n	n	n	n	n	y	n		neg	a		3	no	late	no	
1108	f	10	n	n	n	n	y	y	y	n	n	n	n	n	n	n	n	n		neg	a	y	7	no	no	no	
1109	f	10	y	n	y	y	y	y	n	n	n	n	n	n	n	n	n	n		neg	a		7	no	early	early	
1110	f	10	n	n	y	n	n	n	n	n	n	n	n	n	n	n	y	n		neg	a	n	3	no	no	no	
1111	m	9	y	n	y	n	y	n	n	n	n	n	n	n	n	n	y	n	neg, neg		a	y	7	no	late	no	
1112	m	2	y	n	y	y	y	n	n	n	n	n	n	n	n	n	y	n	PE, neg		a	y	5	early	late	late	
1113	m	10	y	y	y	n	y	y	y	y	y	n	n	n	n	n	n	n	neg, neg		a	n	7	no	cts	cts	
1114	m	9	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	n	neg, neg		a	n	2	no	no	no	
1115	m	10	n	n	y	y	y	y	y	y	y	n	n	n	n	n	n	y	PE, PE		c	y	10	early	early	early	
1116	m	11	n	n	n	n	y	n	y	n	n	n	n	n	n	n	n	n		neg	c	y	5	no	no	no	
1117	f	12	y	y	y	n	y	n	y	n	n	n	n	n	n	n	n	n		neg	a	n	5	no	early	early	
1118	f	1	y	n	y	n	y	y	y	n	y	n	n	n	n	n	n	n		neg	c	n	2	no	early	early	
1120	m	4	n	n	n	n	n	n	n	n	n	y	n	n	n	n	n	n		neg	a	n	4	no	no	no	
1121	m	4	y	y	y	y	y	y	y	n	n	n	n	n	n	n	n	n	neg, neg		a	n	5	no	early	early	
1122	m	11	y	y	y	n	y	y	n	n	n	n	n	n	n	n	n	n		neg	a	n	6	no	early	cts	
1123	m	3	y	n	n	n	y	y	n	n	y	n	n	n	n	n	n	n	prev	neg, neg	a	y	3	no	late	no	
1124	f	4	y	y	y	y	y	y	y	y	y	n	n	n	n	n	n	y	prev		neg	a	y	10	no	late	cts
1125	m	5	n	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n		neg	a	n	4	no	no	no	
1126	m	5	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	neg, neg		a	n	3	no	no	no	
1127	f	5	y	n	y	n	y	n	n	n	n	n	n	n	n	n	n	n	prev	neg, neg	a	n	5	no	early	no	
1128	m	5	n	y	n	y	y	y	y	n	y	n	n	n	n	n	n	n		neg	a	y	5	no	early	early	
1129	m	5	y	y	y	n	y	y	y	n	y	n	n	n	n	n	n	n		neg	a	n	5	no	early	early	
1130	f	5	n	y	y	y	y	y	n	n	y	n	y	n	n	n	y	n		neg	a	y	6	no	late	cts	
1131	m	5	y	n	y	n	y	y	y	n	y	n	n	n	n	n	n	n	neg, neg		a	y	6	no	early	early	
1132	f	6	n	n	y	y	n	y	y	n	n	n	n	n	n	n	n	n		neg	a	y	8.5	no	early	early	
1133	m	5	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	y	neg, neg		a	y	7	no	no	no	
1134	m	5	y	n	y	n	y	y	n	y	y	n	n	n	n	n	y	n		neg	a	n	5	no	cts	cts	
1135	m	6	n	y	y	n	y	n	n	n	n	n	n	n	n	n	n	n		neg	a	n	3	no	early	cts	
1136	f	6	n	n	y	n	n	n	n	y	n	n	n	n	n	n	n	n		neg	a	n	2	no	no	no	
1137	m	6	y	y	y	y	y	y	n	n	y	n	n	n	n	n	n	n		neg	a	y	6	no	early	early	
1138	m	6	y	n	n	n	y	y	n	n	n	n	n	n	n	n	n	n		neg	a	n	9	no	no	no	
1139	f	6	n	y	n	n	y	y	n	y	n	n	n	n	n	n	n	n		neg	a	n	3	no	cts	cts	
1140	f	6	n	y	y	y	y	y	n	n	n	n	n	n	n	n	n	n		neg	c	y	6	no	early	early	
1141	m	8	y	n	y	y	y	n	y	n	y	n	n	n	n	n	n	n		neg	c	n	2	no	early	early	
1142	f	6	n	n	n	n	y	n	y	n	n	n	n	n	n	n	n	n		neg	c	n	3	no	no	no	
1143	m	6	n	n	n	y	y	y	n	n	y	n	n	n	n	n	n	n		neg	a	y	5	no	cts	cts	

Table 21: Patient data from South County Physicians

Wood River Health Services

Expert System Probabilities Without Serum test:					Expert System Probabilities with Serum test:				
Patient #	No Lyme	Early Lyme	Late Lyme	CFS or Fibro.	No Lyme	Early Lyme	Late Lyme	CFS or Fibro.	
501	0.996731	0.000000	0.003244	0.000000	0.999929	0.000000	0.000000	0.000000	
502	0.715840	0.054855	0.000135	0.229171	0.751528	0.007873	0.000000	0.240597	
503	0.998469	0.000831	0.000420	0.000281	0.999597	0.000114	0.000000	0.000281	
504	0.722248	0.042431	0.000958	0.234364	0.750442	0.006027	0.000000	0.243512	
505	0.382945	0.000290	0.000000	0.616749	0.383047	0.000000	0.000000	0.616913	
506	0.995422	0.001498	0.000000	0.003058	0.996733	0.000205	0.000000	0.003062	
507	0.987579	0.003752	0.000858	0.007812	0.985476	0.006672	0.000000	0.007795	
508	0.000250	0.000124	0.000000	0.999625	0.000250	0.000000	0.000000	0.999733	
509	0.001314	0.001140	0.000000	0.997543	0.001315	0.000156	0.000000	0.998529	
510	0.002627	0.000245	0.000000	0.997128	0.002627	0.000000	0.000000	0.997339	
511	0.001295	0.002218	0.000000	0.996481	0.001297	0.000304	0.000000	0.998399	
512	0.008293	0.011807	0.000000	0.979883	0.008378	0.001631	0.000000	0.989991	
513	0.999996	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
514	0.999996	0.000000	0.000000	0.000000	0.999988	0.000000	0.000000	0.000000	
515	0.998753	0.001212	0.000000	0.000000	0.999827	0.000166	0.000000	0.000000	
516	0.999996	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
517	0.005661	0.993762	0.000347	0.000230	0.039939	0.958395	0.000000	0.001620	
518	0.995760	0.002793	0.000101	0.001346	0.955784	0.042691	0.000234	0.001292	
519	0.997738	0.000000	0.002247	0.000000	0.943821	0.000000	0.056171	0.000000	
520	0.547163	0.274343	0.000996	0.177498	0.717889	0.049205	0.000000	0.232881	
521	0.572713	0.239852	0.001593	0.185843	0.723695	0.041432	0.000000	0.234835	
522	0.000245	0.017554	0.001702	0.980499	0.000249	0.002441	0.000000	0.997277	
523	0.999985	0.000000	0.000000	0.000000	0.999938	0.000000	0.000000	0.000000	
524	0.929006	0.042597	0.017830	0.010568	0.982311	0.006157	0.000358	0.011174	
525	0.001297	0.000543	0.000000	0.998157	0.001297	0.000000	0.000000	0.998628	
526	0.999996	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
527	0.999996	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
528	0.999985	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
529	0.970525	0.002398	0.000000	0.027024	0.972589	0.000328	0.000000	0.027082	
530	0.999995	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
531	0.857196	0.009025	0.098565	0.035214	0.957208	0.001378	0.002091	0.039323	
532	0.039853	0.053737	0.000152	0.906258	0.038251	0.091925	0.000000	0.869815	
533	0.333393	0.104606	0.000161	0.561840	0.359722	0.029485	0.004582	0.606211	
534	0.896395	0.077489	0.001156	0.024960	0.961828	0.011366	0.000000	0.026782	
535	0.999993	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
536	0.990864	0.007803	0.000660	0.000673	0.998235	0.001075	0.000000	0.000678	
537	0.997061	0.002204	0.000000	0.000677	0.999019	0.000302	0.000000	0.000678	
538	0.852467	0.145971	0.000395	0.001168	0.268125	0.731209	0.000299	0.000367	
539	0.000253	0.000232	0.000000	0.999514	0.000253	0.000000	0.000000	0.999715	
540	0.992673	0.000000	0.006816	0.000447	0.999411	0.000000	0.000130	0.000450	
541	0.999815	0.000000	0.000000	0.000104	0.999217	0.000571	0.000108	0.000104	
542	0.901785	0.001908	0.053403	0.042905	0.953296	0.000276	0.001073	0.045356	
543	0.999996	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
544	0.999995	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
545	0.999956	0.000000	0.000000	0.000000	0.999995	0.000000	0.000000	0.000000	
546	0.999924	0.000000	0.000000	0.000000	0.999992	0.000000	0.000000	0.000000	
547	0.999984	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
548	0.256351	0.241362	0.001793	0.500493	0.324547	0.041772	0.000000	0.633638	
549	0.007369	0.830945	0.086965	0.074721	0.037345	0.575629	0.008373	0.378653	
550	0.999766	0.000129	0.000105	0.000000	0.997698	0.002049	0.000252	0.000000	
551	0.067768	0.897645	0.034520	0.000000	0.354436	0.641783	0.003430	0.000351	
552	0.999805	0.000117	0.000000	0.000000	0.999982	0.000000	0.000000	0.000000	
553	0.994651	0.002720	0.000422	0.002207	0.997406	0.000373	0.000000	0.002213	
554	0.999037	0.000755	0.000206	0.000000	0.999718	0.000197	0.000000	0.000000	

Table 22: Expert system classifications with and without a serum test



Dr. Joseph England's office

<b>Expert System Probabilities Without Serum test:</b>					<b>Expert System Probabilities with Serum test:</b>				
Patient #	No Lyme	Early Lyme	Late Lyme	CFS or Fibro.	No Lyme	Early Lyme	Late Lyme	CFS or Fibro.	
601	0.989932	0.002911	0.002473	0.004684	0.933253	0.000717	0.061614	0.004416	
602	0.215754	0.776575	0.007192	0.000479	0.668948	0.329145	0.000424	0.001484	
603	0.995524	0.004306	0.000164	0.000000	0.998799	0.001129	0.000000	0.000000	
604	0.017646	0.013300	0.008851	0.960203	0.014520	0.002859	0.192484	0.790137	
605	0.999986	0.000000	0.000000	0.000000	0.999975	0.000000	0.000000	0.000000	
606	0.996037	0.001517	0.000237	0.002210	0.991175	0.000394	0.006232	0.002199	
607	0.962006	0.036336	0.001533	0.000104	0.622723	0.374826	0.002384	0.000000	
608	0.999911	0.000000	0.000000	0.000000	0.999991	0.000000	0.000000	0.000000	
609	0.995957	0.000159	0.000108	0.003775	0.996200	0.003776	0.000000	0.000000	
610	0.997607	0.001692	0.000421	0.000280	0.999480	0.000232	0.000000	0.000000	
611	0.001923	0.997934	0.000112	0.000000	0.007240	0.981485	0.011161	0.000114	
612	0.987336	0.004815	0.000000	0.007808	0.991498	0.000661	0.000000	0.007841	
613	0.999976	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
614	0.999980	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
615	0.976067	0.018233	0.005426	0.000274	0.997069	0.002546	0.000105	0.000280	
616	0.998832	0.000466	0.000000	0.000678	0.999258	0.000000	0.000000	0.000678	
617	0.999030	0.000374	0.000575	0.000000	0.999917	0.000000	0.000000	0.000000	
618	0.978461	0.020225	0.000652	0.000662	0.996498	0.002816	0.000000	0.000674	
619	0.999946	0.000000	0.000000	0.000000	0.999971	0.000000	0.000000	0.000000	
620	0.999271	0.000000	0.000000	0.000654	0.999344	0.000000	0.000000	0.000654	
622	0.830831	0.048656	0.000333	0.120180	0.481163	0.448773	0.000463	0.069601	
623	0.000216	0.000000	0.000000	0.999780	0.000000	0.000000	0.000000	0.999967	
624	0.615210	0.000136	0.000000	0.384648	0.615286	0.000000	0.000000	0.384695	
625	0.999984	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
626	0.991673	0.008282	0.000000	0.000000	0.998852	0.001140	0.000000	0.000000	
627	0.999434	0.000490	0.000000	0.000000	0.997883	0.000128	0.001989	0.000000	
628	0.997953	0.000000	0.002043	0.000000	0.999961	0.000000	0.000000	0.000000	
629	0.000000	0.000000	0.000000	0.999974	0.000000	0.000000	0.000000	0.999980	

Table 22 (cont.): Expert system classification with and without a serum test

South County Physicians

<i>Expert System Probabilities Without Serum test:</i>					<i>Expert System Probabilities with Serum test:</i>				
Patient #	No Lyme	Early Lyme	Late Lyme	CFS or Fibro.	No Lyme	Early Lyme	Late Lyme	CFS or Fibro.	
1101	0.998197	0.000416	0.000400	0.000987	0.998947	0.000000	0.000000	0.000988	
1102	0.999616	0.000000	0.000000	0.000325	0.999673	0.000000	0.000000	0.000325	
1103	0.009537	0.061060	0.000000	0.929330	0.010069	0.008812	0.000000	0.981118	
1104	0.982684	0.009461	0.000000	0.007773	0.990856	0.001304	0.000000	0.007838	
1105	0.998098	0.000957	0.000891	0.000000	0.999798	0.000131	0.000000	0.000000	
1106	0.999968	0.000000	0.000000	0.000000	0.999994	0.000000	0.000000	0.000000	
1107	0.946286	0.004614	0.041616	0.007483	0.990678	0.000660	0.000828	0.007834	
1108	0.998098	0.000957	0.000891	0.000000	0.999798	0.000131	0.000000	0.000000	
1109	0.863223	0.100943	0.000363	0.035471	0.945998	0.015122	0.000000	0.038872	
1110	0.997045	0.000000	0.002934	0.000000	0.999936	0.000000	0.000000	0.000000	
1111	0.971492	0.000574	0.027274	0.000659	0.998709	0.000000	0.000533	0.000678	
1112	0.868831	0.002188	0.125917	0.003064	0.718490	0.028816	0.250160	0.002534	
1113	0.001295	0.002218	0.000000	0.996481	0.001297	0.000304	0.000000	0.998399	
1114	0.999952	0.000000	0.000000	0.000000	0.999989	0.000000	0.000000	0.000000	
1115	0.000142	0.041351	0.000000	0.958496	0.000000	0.407220	0.000000	0.592674	
1116	0.999394	0.000000	0.000535	0.000000	0.999976	0.000000	0.000000	0.000000	
1117	0.573876	0.066433	0.000777	0.358913	0.609284	0.009642	0.000000	0.381056	
1118	0.593825	0.079570	0.001667	0.324939	0.638746	0.011700	0.000000	0.344520	
1120	0.999988	0.000000	0.000000	0.000000	0.999990	0.000000	0.000000	0.000000	
1121	0.041012	0.888933	0.000921	0.069134	0.177019	0.524503	0.000000	0.298403	
1122	0.739132	0.020912	0.000113	0.239843	0.752806	0.002912	0.000000	0.244280	
1123	0.965342	0.000165	0.030833	0.003660	0.995598	0.000000	0.000604	0.003775	
1124	0.000000	0.000651	0.051677	0.947669	0.000000	0.000000	0.001035	0.998668	
1125	0.999952	0.000000	0.000000	0.000000	0.999977	0.000000	0.000000	0.000000	
1126	0.999993	0.000000	0.000000	0.000000	0.999999	0.000000	0.000000	0.000000	
1127	0.969669	0.010073	0.019600	0.000658	0.997523	0.001416	0.000383	0.000677	
1128	0.440536	0.189855	0.013762	0.355847	0.535543	0.031550	0.000318	0.413289	
1129	0.031717	0.246497	0.000348	0.721438	0.040309	0.042824	0.000000	0.916859	
1130	0.007370	0.012588	0.109177	0.870864	0.008356	0.001951	0.002352	0.987341	
1131	0.361693	0.436272	0.001488	0.200547	0.581587	0.095896	0.000000	0.322471	
1132	0.666067	0.332848	0.000826	0.000259	0.993570	0.063919	0.000000	0.000364	
1133	0.993946	0.005217	0.000836	0.000000	0.999267	0.000717	0.000000	0.000000	
1134	0.000760	0.000000	0.000139	0.999075	0.000760	0.000000	0.000000	0.999234	
1135	0.950963	0.022427	0.000130	0.026479	0.969865	0.003127	0.000000	0.027006	
1136	0.995934	0.003919	0.000140	0.000000	0.999453	0.000538	0.000000	0.000000	
1137	0.007672	0.085731	0.000000	0.906562	0.008286	0.012657	0.000000	0.979057	
1138	0.999363	0.000424	0.000159	0.000000	0.999885	0.000000	0.000000	0.000000	
1139	0.002703	0.000000	0.000000	0.997281	0.002703	0.000000	0.000000	0.997295	
1140	0.176690	0.529201	0.000262	0.293846	0.325466	0.133255	0.000000	0.541269	
1141	0.115920	0.852175	0.003623	0.028283	0.444538	0.446736	0.000264	0.108461	
1142	0.986909	0.000899	0.001469	0.010723	0.989102	0.000123	0.000000	0.010747	
1143	0.973962	0.006088	0.001024	0.018926	0.980098	0.000837	0.000000	0.019045	

Table 22 (cont.): Expert system classifications with and without a serum test

	<i>Fisher's Exact Test of independence for patients with Lyme</i>								
	<i>Myalgia</i>	<i>Arthralgia</i>	<i>Fever</i>	<i>Fatigue</i>	<i>Tender Pts.</i>	<i>Sleep Dis</i>	<i>Headache</i>	<i>Stiff neck</i>	<i>Arthritis</i>
<i>Myalgia</i>	-		significant						
<i>Arthralgia</i>		-	significant						
<i>Fever</i>	significant	significant	-	significant					
<i>Fatigue</i>			significant	-					
<i>Tender Pts</i>					-				
<i>Sleep Dis.</i>						-			
<i>Headache</i>							-		
<i>Stiff neck</i>								-	
<i>Athritis</i>									-

Table 23: Results from Fisher's exact test for independence.

Misclassified patients, patients that tested positive for Lyme but were classified by the expert system as not having Lyme																										
Patient #	Sex	Onset	Recall bite	Myalgia	Arthralgia	Fever	Fat. or Mal.	Headache	Stiff neck	Tender Pts.	Sleep dis.	Bells palsy	radiculo.	Meningitis	Enceph.	Nodal	Arthritis	ECM	2 serum	1 serum	Age	Treated	Likelihood	TRUE	Expert System	
514	f	11	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	PE,PE	a	n	4	early	no	no	
515	f	11	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	y	neg,PE	a	y	10	early	no	no	
518	m	10	y	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	PE,PE	c	n	5	early	no	no	
523	f	12	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	PE,PE	c	n	2	early	no	no	
526	m	3	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	XR,PE	a	n	2	early	no	no	
541	m	5	y	n	n	n	n	n	n	y	n	n	n	n	n	n	n	y	PE,PE	a	y	8	early	no	no	
546	f	5	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	y	neg,PE	a	y	9	early	no	no	
550	f	7	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y	PE	c	y		early	no	no	
619		10	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y	XR,PE			y	10	early	no	no
519		11	n	n	y	n	n	n	n	n	n	n	n	n	n	n	y	n	PL,PL	a	y	7	late	no	no	
601	f	6	y	y	n	n	n	y	n	n	n	y	n	n	n	n	n	n	PL,PL	a	y	7	late	no	no	
605	f	8	n	n	n	n	n	n	n	n	n	y	n	n	n	n	n	n	PE,PE	a	y	4	late	no	no	
606		7	n	y	n	n	n	y	n	n	n	n	n	n	n	n	n	n	PL	a	y	4	late	no	no	
627	f	6	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	y	PL,PL	c	y	7	late	no	no	
Misclassified patients, patients that tested negative for Lyme but were classified by the expert system as having Lyme																										
Patient #	Sex	Onset	Recall bite	Myalgia	Arthralgia	Fever	Fat. or Mal.	Headache	Stiff neck	Tender Pts.	Sleep dis.	Bells palsy	radiculo.	Meningitis	Enceph.	Nodal	Arthritis	ECM	2 serum	1 serum	Age	Treated	Likelihood	TRUE	Expert System	
502	f	8	n	y	y	n	y	y	n	n	n	n	n	n	n	n	n	n	neg, neg	neg	c		6	no	early	
504	f	5	y	y	y	n	y	y	y	n	n	n	y	n	n	n	n	n	neg, neg	a	y		4	no	early	
517	f	11	y	n	y	y	y	y	y	n	n	n	n	n	n	n	n	y	neg, neg	c	y		7	no	early	
520	n	11	n	y	y	n	y	y	n	n	n	n	n	n	n	n	n	n	neg, neg	a	n		4	no	early	
521	f	12	y	y	y	n	y	y	y	n	n	n	n	n	n	n	n	n	neg, neg	a	y		5	no	early	
522	f	11	y	y	y	y	y	y	n	y	n	n	n	n	n	n	n	n	prev	neg, neg	a	y		6	no	early
524	f	12	y	y	n	y	y	y	y	n	n	n	n	n	n	n	n	n	neg	neg	c	y		8	no	early
532	f	4	n	y	y	n	y	y	n	y	n	n	n	n	n	n	n	n	XR, XR	a	n		3	no	early	
534	m	4	n	y	y	n	y	n	y	n	n	n	n	n	n	n	n	n	neg, neg	a	n		3	no	early	
548	m	5	y	y	y	n	y	n	y	n	y	n	n	n	n	n	n	n	neg, neg	a	y		5	no	early	
549	m	6	n	y	y	y	y	n	y	n	y	n	n	n	n	n	n	n	prev	neg	a		4	no	early	
551	m	6	y	y	n	y	y	n	y	n	n	n	n	n	n	n	n	n	neg	neg	a	y		8	no	early
602	m	6	n	y	n	n	y	y	y	n	n	n	n	n	n	n	n	y	neg, neg	a	y		7	no	early	
615	m	9	n	n	n	y	y	y	n	n	n	n	n	n	n	n	n	n	neg	neg	a	n		4	no	early
618		9	n	n	y	n	y	n	y	n	n	n	n	n	n	n	n	n	neg	neg	n		2	no	early	
1103	f	8	n	n	y	y	y	y	y	n	n	n	n	n	n	n	n	n	neg, neg	a	y		3	no	early	
1109	f	10	y	n	y	y	y	y	n	n	n	n	n	n	n	n	n	n	neg	neg	a		7	no	early	
1117	f	12	y	y	y	n	n	n	y	y	n	n	n	n	n	n	n	n	neg	neg	a	n		5	no	early
1118	f	1	y	n	y	n	y	y	y	n	y	n	n	n	n	n	n	n	neg	neg	c	n		2	no	early
1121	m	4	y	y	y	y	y	y	y	n	n	n	n	n	n	n	n	n	neg, neg	a	n		5	no	early	
1122	m	11	y	y	y	n	y	y	n	n	n	n	n	n	n	n	n	n	neg	neg	a	n		6	no	early
1127	f	5	y	n	y	n	y	n	n	n	n	n	n	n	n	n	n	n	prev	neg, neg	a	n		5	no	early
1128	m	5	n	y	n	y	y	y	y	n	y	n	n	n	n	n	n	n	neg	neg	a	y		5	no	early
1129	m	5	y	y	y	n	y	y	n	y	n	n	n	n	n	n	n	n	neg	neg	a	n		5	no	early
1131	m	5	y	n	y	n	y	y	y	n	y	n	n	n	n	n	n	n	neg, neg	a	y		6	no	early	
1132	f	6	n	n	y	y	n	y	y	n	n	n	n	n	n	n	n	n	neg	neg	a	y		8.5	no	early
1135	m	6	n	y	y	n	y	n	n	n	n	n	n	n	n	n	n	n	neg	neg	a	n		3	no	early
1137	m	6	y	y	y	y	y	y	n	n	y	n	n	n	n	n	n	n	neg	neg	a	y		6	no	early
1140	f	6	n	y	y	y	y	y	n	n	n	n	n	n	n	n	n	n	neg	neg	c	y		6	no	early
1141	m	6	y	n	y	y	y	n	y	n	y	n	n	n	n	n	n	n	neg	neg	c			2	no	early
531	f	2	n	n	y	y	y	y	n	n	n	n	n	n	n	n	n	y	neg, neg	a	y		9	no	late	
542	m	5	n	n	y	n	y	n	n	n	y	n	n	n	n	n	n	y	neg, neg	a	n		5	no	late	
1107	f	9	n	n	y	n	n	y	n	n	n	n	n	n	n	n	n	y	neg	neg	a			3	no	late
1111	m	9	y	n	y	n	y	n	n	n	n	n	n	n	n	n	n	y	neg, neg	a	y		7	no	late	
1123	m	3	y	n	n	n	y	y	n	n	y	n	n	n	n	n	n	n	prev	neg, neg	a	y		3	no	late
1124	f	4	y	y	y	y	y	y	y	y	y	n	n	n	n	n	n	y	prev	neg	a	y		10	no	late
1130	f	5	n	y	y	y	y	y	n	n	y	n	y	n	n	n	n	y	neg	neg	a	y		6	no	late

Table 24: A table of the misclassified patients

Patient #	Sex	Onset	Recall bite	Myalgia	Arthralgia	Fever	Fat. or Mal.	Headache	Stiff neck	Tender Pts.	Sleep dis.	Bells palsy	radiculo.	Meningitis	Enceph.	Nodal	Arthritis	ECM	2 serum	1 serum	Age	Treated	Likelihood	TRUE	Expert	Expert	
514	f	11	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	PE,PE		a	n	4	early	no	no	
515	f	11	y	n	y	n	n	n	n	n	n	n	n	n	n	n	n	y	neg,PE		a	y	10	early	no	no	
518	m	10	y	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	PE,PE		c	n	5	early	no	early	
523	f	12	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	PE,PE		c	n	2	early	no	no	
526	m	3	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	n	XR,PE		a	n	2	early	no	no	
538	m	4	y	y	y	y	n	n	n	n	n	n	n	n	n	n	n	y		PE	a	y	10	early	early	early	
541	m	5	y	n	n	n	n	n	n	y	n	n	n	n	n	n	n	y	PE,PE		a	y	8	early	no	no	
546	f	5	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	y	neg,PE		a	y	9	early	no	no	
550	f	7	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y		PE	c	y		early	no	no	
607	f	8	y	y	n	y	n	y	n	n	n	n	n	n	n	n	n	y	PE,PE		c	y	10	early	early	early	
619		10	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n	y	XR,PE			y	10	early	no	no	
622	m	9	n	y	y	y	y	n	n	n	n	n	n	n	n	n	n	n	PE,rem		a	y	3	early	early	early	
1112	m	2	y	n	y	y	y	n	n	n	n	n	n	n	n	n	y	n	PE,neg		a	y	5	early	late	late	
1115	m	10	n	n	y	y	y	y	y	y	y	n	n	n	n	n	n	y	PE,PE		c	y	10	early	early	early	
519		11	n	n	y	n	n	n	n	n	n	n	n	n	n	n	n	y	n	PL,PL		a	y	7	late	no	late
533	m	4	n	y	y	y	y	y	n	n	n	n	n	n	n	n	n	n	PL,PL		a	y	7	late	early	early	
601	f	6	y	y	n	n	n	y	y	n	n	y	n	n	n	n	n	n	PL,PL		a	y	7	late	no	late	
604	m	7	n	y	y	y	y	n	n	n	y	n	n	n	n	n	y	n		PL	c			late	early	late	
605	f	8	n	n	n	n	n	n	n	n	n	y	n	n	n	n	n	n	PE,PE		a	y	4	late	no	no	
606		7	n	y	n	n	y	y	n	n	n	n	n	n	n	n	n	n		PL	a	y	4	late	no	no	
611		7	y	y	y	y	n	y	y	n	n	n	n	n	n	n	n	y	PL,PL		c	y	6	late	early	early	
627	f	6	n	n	n	n	n	y	n	n	n	n	n	n	n	n	n	y	PL,PL		c	y	7	late	no	no	

Table 25: A table of early and late Lyme patients

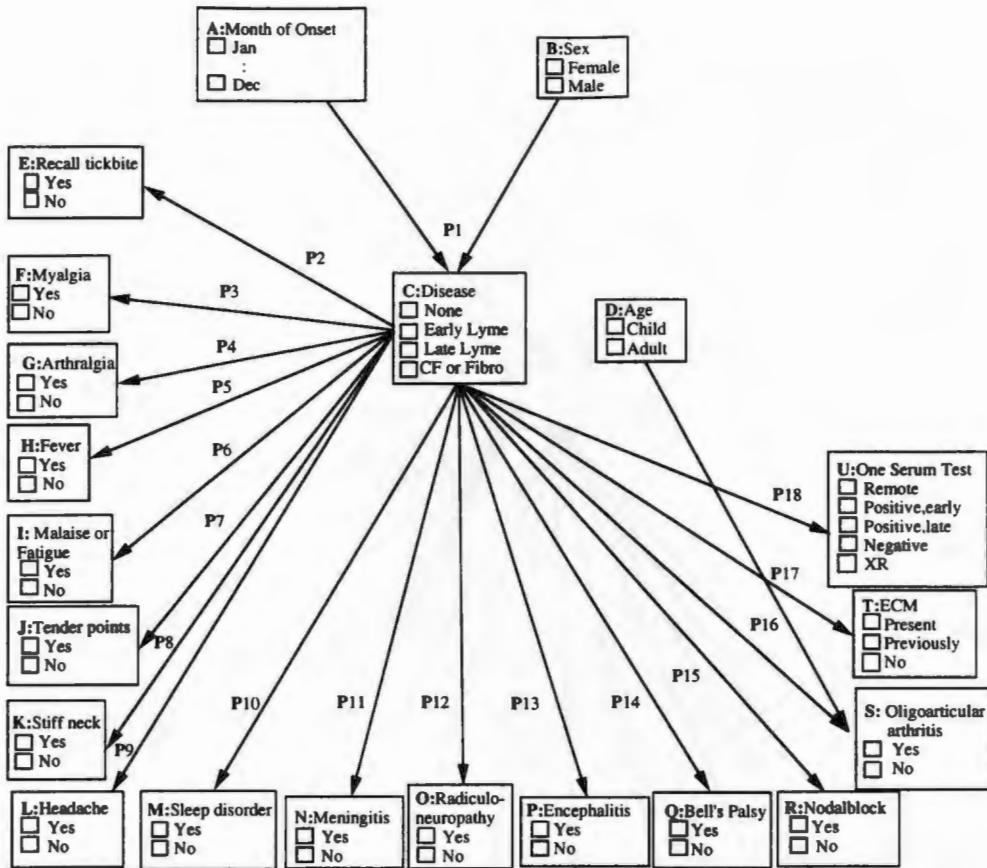
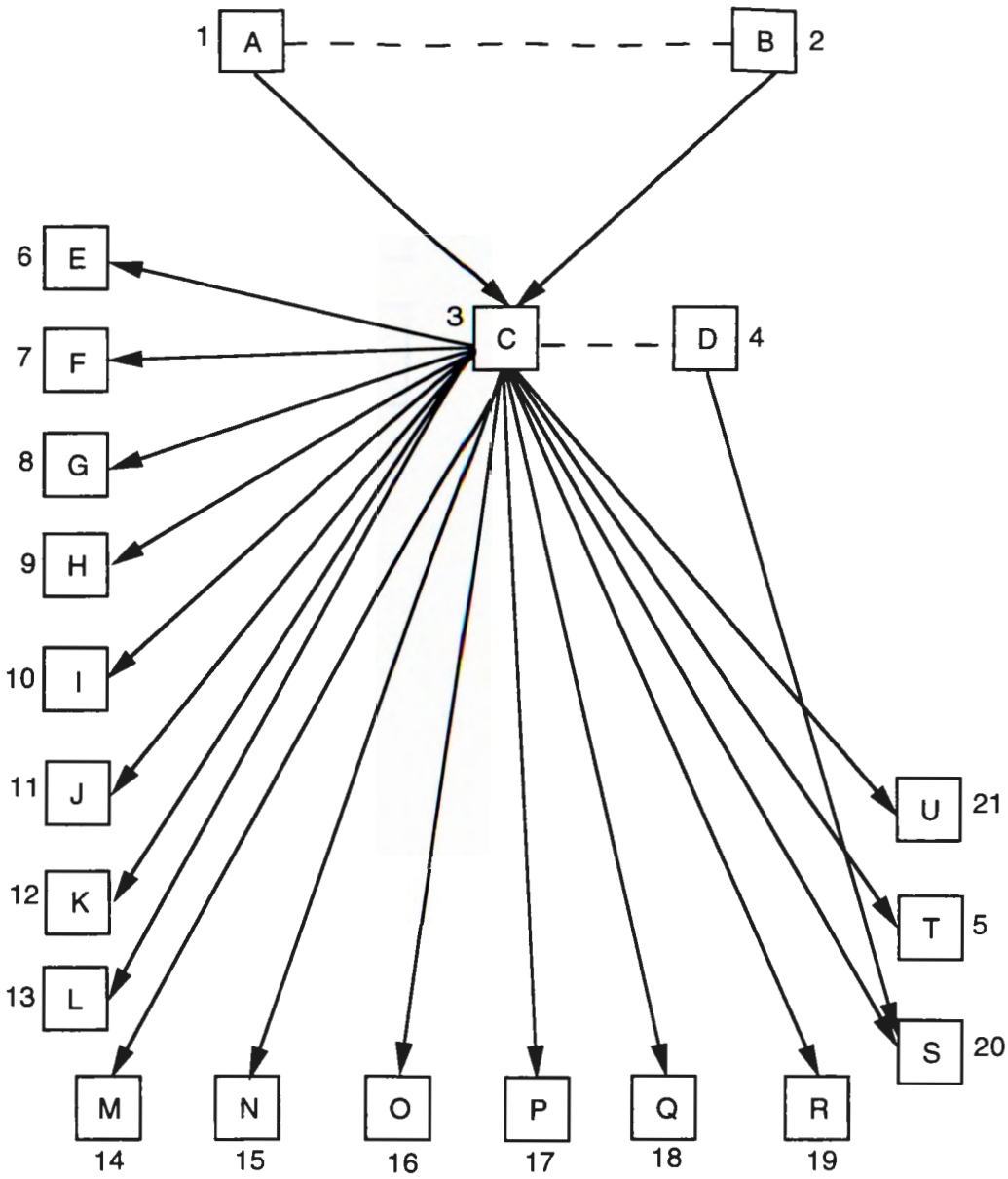


Figure 1: The graph of nodes before triangulation



- - - - A link that has been added to triangulate the graph.  
 \_\_\_\_\_ An original link from the tree of nodes.

Figure 2: The graph of nodes after triangulation.

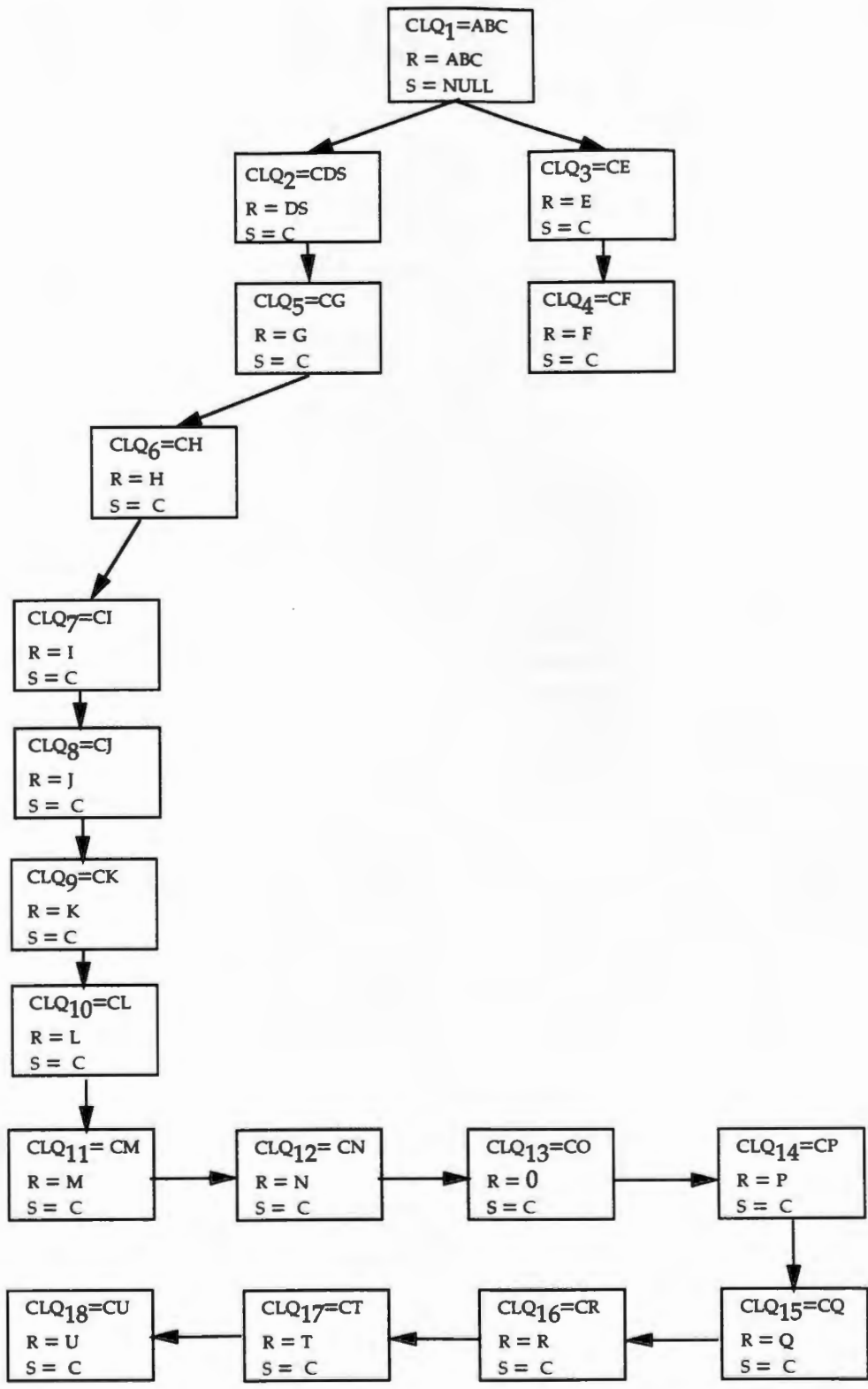


Figure 3: The tree of cliques.



Clq no.	Parent	Nodes	$R_i$	$S_i$	Configuration	$\Psi(\text{Clq}_i)$	$P(\text{Clq}_i)$
Clq <sub>1</sub>	---	{ A,B,C }	{ A,B,C }	{ null }	{ a1,b1, c1 } { a1, b1, c2 }  ⋮ { a12, b2, c4 }		
Clq <sub>2</sub>	1	{ C,D,S }	{ D,S }	{ C }	{ c1,d1,s1 } { c1,d1,s2 } { c1,d2,s1 } { c1,d2,s2 }  ⋮ { c4,d2,s1 } { c4,d2,s2 }		
Clq <sub>3</sub>	1	{ C, E }	{ E }	{ C }	{ c1, e1 } { c1, e2 } { c2, e1 } { c2, e2 } { c3, e1 } { c3, e2 } { c4, e1 } { c4, e2 }		
Clq <sub>4</sub>	2	{ C, F }	{ F }	{ C }	{ c1, f1 } { c1, f2 } { c2, f1 } { c2, f2 } { c3, f1 } { c3, f2 } { c4, f1 } { c4, f2 }		

Figure 4: The partial list of configurations inside a particular clique.

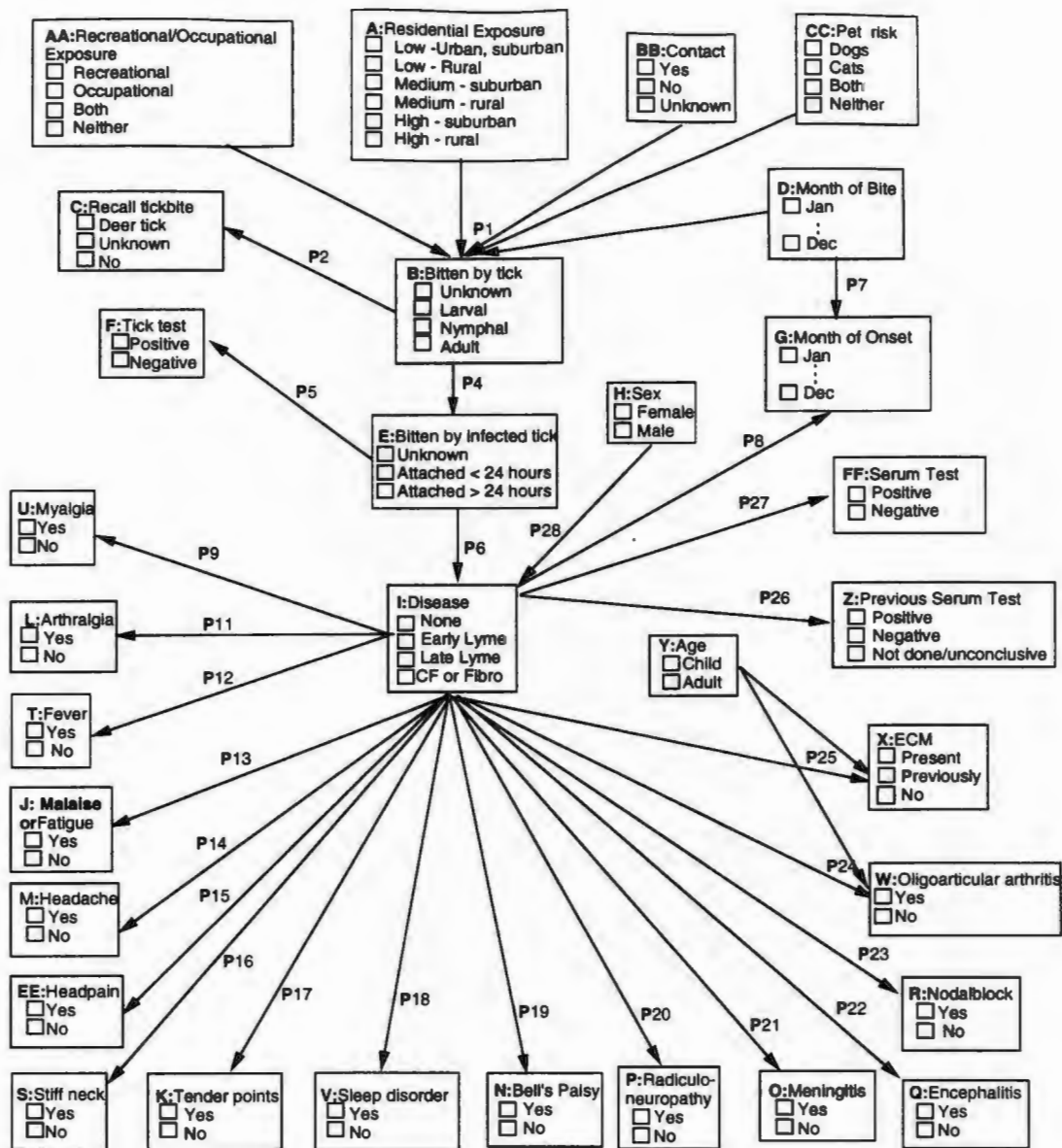
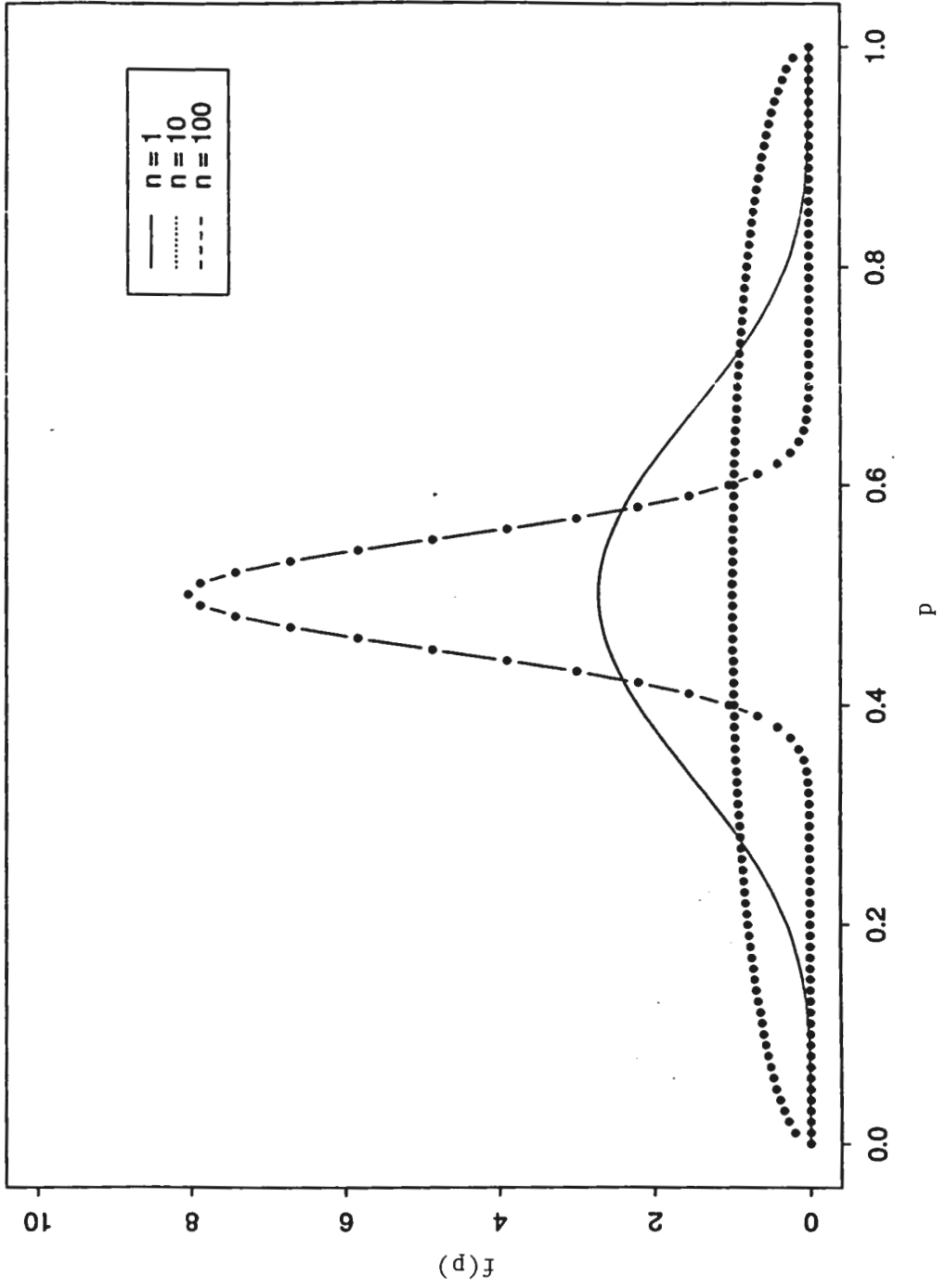


Figure 5: A graph of nodes with the addition of diagnostic historical variables.

Figure 6: A Plot of the Beta Distribution



Scatter Plot Likelihood vs. Expert

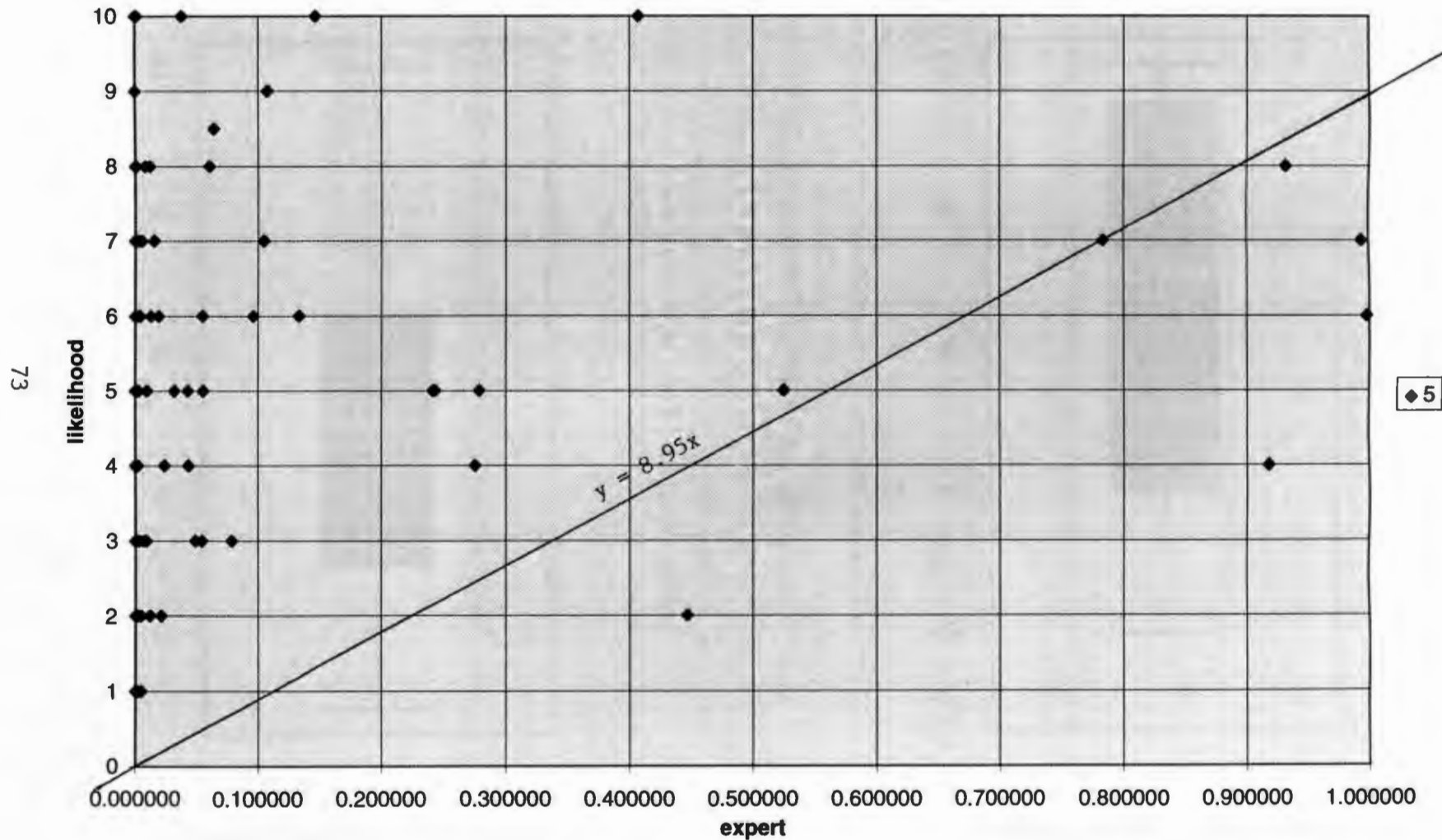
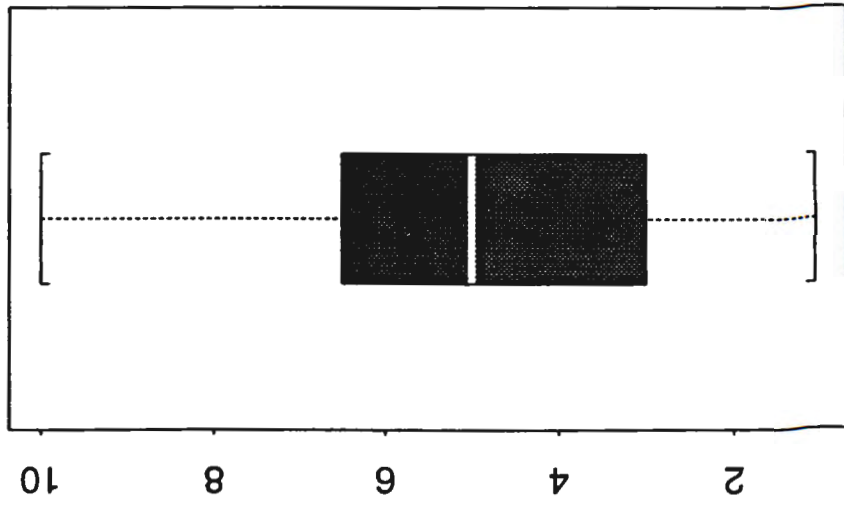
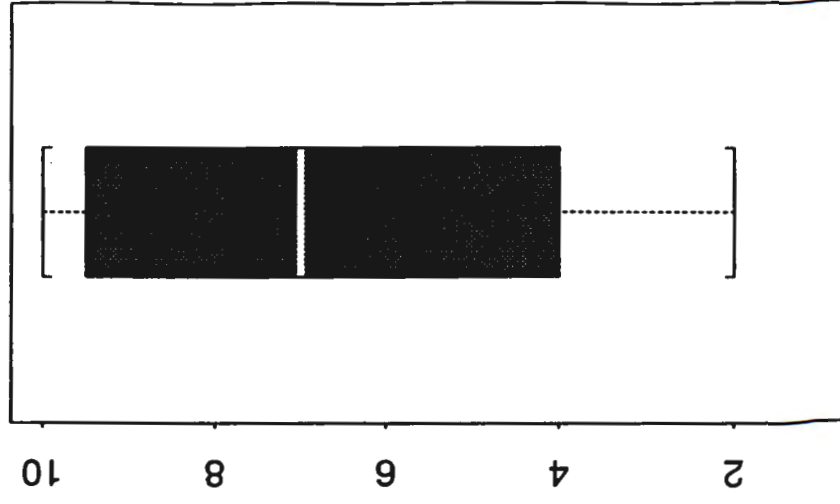


Figure 7: A scatter plot of physicians likelihood of Lyme disease vs. expert system's probabilities without serum test

Figure 8: Boxplots of Physicians' Likelihood



Patients without Lyme Disease



Patients with Lyme Disease

## References

1. Anderson, J.F., R.C. Johnson, L.A. Magnarelli. 1987. Seasonal prevalence of *Borrelia burgdorferi* in natural populations of white-footed mice, *Peromyscus leucopus*. *Journal of Clinical Microbiology* 25(8), 1564-1566.
2. Berger, B.W., R.C. Johnson, C. Kodner, L. Coleman. 1995. Cultivation of *Borrelia burgdorferi* from human tick bite sites: a guide to the risk of infection. *Journal of the American Academy of Dermatology* 32, 184-187.
3. Blaauw AAM, MKE Nohlmans, AEJM van den Bogaard, S van der Linden. 1992. Diagnostic tools in Lyme borreliosis: clinical history compared with serology. *Journal of Clinical Epidemiology* II 1229-1236.
4. Bowen, G.S., T.L. Schulze, W.L. Parkin. 1984. Lyme disease in New Jersey, 1978-1982. *The Yale Journal of Biology and Medicine* 57, 661-668.
5. Buchwald, D., D. Garrity. 1994. Comparison of patients with chronic fatigue syndrome, fibromyalgia, and multiple chemical sensitivities. *Archives of Internal Medicine*. 154, 2049-2053.
6. Costello, C.M., A.C. Steere, R.E. Pinkerton, H.M. Feder. 1989. A prospective study of tick bites in an endemic area for Lyme disease. *Journal of Infectious Diseases* 159, 136-139.
7. Dekonenko, E.P. 1987. Clinical manifestations of tick-borne erythema in the USSR. *Proceedings of the Second International Symposium on Lyme Disease and Related Disorders*, Vienna, 1985. Stuttgart; New York: 452, 453.
8. Dinerman, H., A.C. Steere. 1992. Lyme disease associated with fibromyalgia. *Annals of Internal Medicine* 117, 281-285.
9. Eichenfield, A.H., D.P. Goldsmith, J.L. Benach, et al. 1986. Lyme arthritis: experience in an endemic area. *Journal of Pediatrics* 109, 753-758.

10. Falco, R.C., D. Fish. 1988.  
Ticks parasitizing humans in a Lyme disease endemic area of southern New York state.  
*American Journal of Epidemiology* 128, 1146-1152.
11. Goldenberg, D.L. 1987.  
Fibromyalgia syndrome: an emerging but controversial condition.  
*JAMA* 257(20), 2782-2787.
12. Goldenberg D.L. 1988.  
Fibromyalgia and other chronic fatigue syndromes: is there evidence for chronic viral disease?  
*Seminars in Arthritis and Rheumatism* 18, 111-120.
13. Goldenberg, D.L., R.W. Simms, A. Geiger, A.L. Komaroff. 1990.  
High frequency of fibromyalgia in patients with chronic fatigue seen in a primary care practice.  
*Arthritis and Rheumatism* 33(3), 381-387.
14. Halperin, J.J. M. Golightly. 1992.  
Lyme borreliosis in Bell's palsy.  
*Neurology* 42, 1268-1270.
15. Hanrahan, J.P., J.L. Benach, J.L. Coleman, E.M. Bosler, J.C. Grabau, D.L. Morse. 1984.  
Epidemiologic features of Lyme disease in New York.  
*The Yale Journal of Biology and Medicine* 57, 643-650.
16. Hsu, V.M., S.J. Patella, L.H. Sigal. 1993.  
"Chronic Lyme disease" as the incorrect diagnosis in patients with fibromyalgia.  
*Arthritis and Rheumatism* 36(11), 1493-1500.
17. Huppertz, H.I. 1990.  
Childhood Lyme borreliosis in Europe.  
*European Journal Pediatrics* 149, 814-821.
18. Kaplan, R.F., M.E. Meadows, L.C. Vincent, E.L. Logigian, A.C. Steere. 1992.  
Memory impairment and depression in patients with Lyme encephalopathy: comparison with fibromyalgia and nonpsychotically depressed patients.  
*Neurology* 42, 1263-1267.

19. Magid, D., B. Schwartz, J. Craft, J.S. Schwartz. 1992. Prevention of Lyme disease after tick bites: a cost effectiveness analysis. *The New England Journal of Medicine* 327(8), 534-540.
20. Magnarelli, L.A., J.F. Anderson, C.S. Apperson, D. Fish, R.C. Johnson, W.A. Chappell. 1986. Spirochetes in ticks and antibodies to *Borrelia burgdorferi* in white-tailed deer from Connecticut, New York state, and North Carolina. *Journal of Wildlife Diseases* 22(2), 178-188.
21. Neapolitan R.E. 1990. *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*, John Wiley and Sons, Inc., New York.
22. Petersen, L.R., A.H. Sweeney, P.J. Checko, L.A. Magnarelli, P.A. Mshar, R.A. Gunn, J.L. Hadler. 1989. Epidemiological and clinical features of 1149 persons with Lyme disease identified by laboratory-based surveillance in Connecticut. *The Yale Journal of Biology and Medicine* 62, 253-262.
23. Peisman, J., A. Spielman. 1979. Host-associations and seasonal abundance of immature *Ixodes dammini* in southeastern Massachusetts. *Annals of the Entomological Society of America* 72(6), 829-831.
24. Piesman, J., T.N. Mather, R.J. Sinsky, A. Spielman. 1987. Duration of tick attachment and *Borrelia burgdorferi* transmission. *Journal of Clinical Microbiology* 25(3), 557-558.
25. Piesman, J., T.N. Mather, G.J. Dammin, S.R. Telford III, C.C. Lastavica, A. Spielman. 1987. Seasonal variation of transmission risk of Lyme disease and human Babesiosis. *American Journal of Epidemiology* 126(6), 1187-1189.
26. Piesman, J. 1993. Dynamics of *Borrelia burgdorferi* transmission by nymphal *Ixodes dammini* ticks. *Journal of Infectious Diseases* 167, 1082-1085.
27. Rees, D.H.E., J.S. Axford. 1994. Evidence for Lyme disease in urban park workers: a potential new health hazard for city inhabitants. *British Journal of Rheumatology* 33, 123-128.



28. Schulze, T.L., G.S. Bowen, M.F. Lakat, W.E. Parkin, J.K. Shisler. 1986  
Seasonal abundance and hosts of *Ixodes dammini* (Acari: Ixodidae)  
and other Ixodid ticks from an endemic Lyme disease focus in New Jersey,  
USA.  
*Journal of Medical Entomology* 23(1), 105-109.
29. Shih, C-M., A. Spielman. 1993.  
Accelerated transmission of Lyme disease spirochetes by partially fed  
vector ticks.  
*Journal of Clinical Microbiology* 31, 2878-2881.
30. Sigal, L.H. 1990.  
Summary of the first 100 patients seen at a Lyme disease referral center.  
*The American Journal of Medicine* 88, 577-581.
31. Sigal, L.H., S.J. Patella. 1992.  
Lyme arthritis as the incorrect diagnosis in pediatric and adolescent  
fibromyalgia.  
*Pediatrics* 90(4), 523-528.
32. Spiegelhalter, D.J., Dawid, A.P., Lauritzen, S.L., Cowell, R.G. 1993.  
Bayesian Analysis in Expert Systems.  
*Statistical Science* 8(3), 219-233.
33. Stafford, K.C. III, L.A. Magnarelli. 1993.  
Spatial and temporal patterns of *Ixodes scapularis* (Acari: Ixodidae)  
in southeastern Connecticut.  
*Journal of Medical Entomology* 30(4), 762-771.
34. Steere A.C., S.E. Malawista, J.A. Hardin, S. Ruddy, P.W. Askenase,  
W.A. Andiman. 1977.  
Erythema chronicum migrans and Lyme Arthritis.  
*Annals of Internal Medicine* 86(6), 685-698.
35. Steere, A.C., A. Gibofsky, M.E. Patarroyo, R.J. Winchester, J.A. Hardin,  
S.E. Malawista. 1979.  
Chronic Lyme arthritis.  
*Annals of Internal Medicine* 90, 896-901.
36. Steere, A.C., N.H. Bartenhagen, J.E. Craft, G.J. Hutchinson, J.H. Newman,  
D.W. Rahn, L.H. Sigal, P.N. Spieler, K.S. Stenn, S.E. Malawista. 1983.  
The early clinical manifestations of Lyme disease.  
*Annals of Internal Medicine* 99, 76-82.

37. Steere, A.C., E. Taylor, M.L. Wilson, J.F. Levine, A. Spielman. 1986.  
Longitudinal assessment of the clinical and epidemiological features  
of Lyme disease in a defined population.  
*The Journal of Infectious Diseases* 154(2), 295-300.
38. Steere, A.C., E. Taylor, G.L. McHugh, E.L. Logigian. 1993.  
The overdiagnosis of Lyme disease.  
*JAMA* 269(14), 1812-1816.
39. Wahlberg, P., H. Granlund, D. Nyman, J. Panelius, I. Seppala. 1993.  
Late Lyme borreliosis: epidemiology, diagnosis and clinical features.  
*Annals of Medicine* 25, 349-352.
40. Waylonis, G.W., W. Heck. 1992.  
Fibromyalgia syndrome.  
*American Journal of Physical Medicine and Rehabilitation* 71(6), 343-348.
41. Whelton, C.L., I. Salit, H. Moldofsky. 1992.  
Sleep, Epstein-Barr virus infection, musculoskeletal pain, and depressive  
symptoms in chronic fatigue syndrome.  
*The Journal of Rheumatology* 19(6), 939-943.

## **Bibliography**

- Anderson, J.F., R.C. Johnson, L.A. Magnarelli. 1987.  
Seasonal prevalence of *Borrelia burgdorferi* in natural  
populations of white-footed mice, *Peromyscus leucopus*.  
*Journal of Clinical Microbiology* 25(8), 1564-1566.
- Barbour, A.G., and Fish D. 1993.  
The biological and social phenomenon of Lyme disease.  
*Science*, Vol. 260: 1610-1615.
- Berger, B.W., R.C. Johnson, C. Kodner, L. Coleman. 1995.  
Cultivation of *Borrelia burgdorferi* from human tick  
bite sites: a guide to the risk of infection.  
*Journal of the American Academy of Dermatology* 32, 184-187.
- Blaauw AAM, MKE Nohlmans, AEJM van den Bogaard, S van der Linden. 1992.  
Diagnostic tools in Lyme borreliosis: clinical history compared with serology.  
*Journal of Clinical Epidemiology II* 1229-1236.
- Blaauw, A.A.M. 1987.  
PhD thesis, University of the Netherlands
- Bowen, G.S., T.L. Schulze, W.L. Parkin. 1984.  
Lyme disease in New Jersey, 1978-1982.  
*The Yale Journal of Biology and Medicine* 57, 661-668.
- Buchwald, D., D. Garrity. 1994.  
Comparison of patients with chronic fatigue syndrome, fibromyalgia, and  
multiple chemical sensitivities.  
*Archives of Internal Medicine*. 154, 2049-2053.
- CDC, Case definitions for public health surveillance. 1994.  
*Morbidity and Mortality Weekly Report*, Vol. 43 (31): 564-571.
- Costello, C.M., A.C. Steere, R.E. Pinkerton, H.M. Feder. 1989.  
A prospective study of tick bites in an endemic area for Lyme disease.  
*Journal of Infectious Diseases* 159, 136-139.
- Dennis, D.T. 1991.  
Lyme disease, tracking an epidemic.  
*Journal of the American Medical Association* Vol. 266: 1269-1270.

- Dekonenko, E.P. 1987.  
Clinical manifestations of tick-borne erythema in the USSR.  
*Proceedings of the Second International Symposium on Lyme Disease and Related Disorders*, Vienna, 1985. Stuttgart; New York: 452, 453.
- Dinerman, H., A.C. Steere. 1992.  
Lyme disease associated with fibromyalgia.  
*Annals of Internal Medicine* 117, 281-285.
- Eichenfield, A.H., D.P. Goldsmith, J.L. Benach, et al. 1986.  
Lyme arthritis: experience in an endemic area.  
*Journal of Pediatrics* 109, 753-758.
- Falco, R.C., D. Fish. 1988.  
Ticks parasitizing humans in a Lyme disease endemic area of southern New York state.  
*American Journal of Epidemiology* 128, 1146-1152.
- Goldenberg, D.L. 1987.  
Fibromyalgia syndrome: an emerging but controversial condition.  
*JAMA* 257(20), 2782-2787.
- Goldenberg D.L. 1988.  
Fibromyalgia and other chronic fatigue syndromes: is there evidence for chronic viral disease?  
*Seminars in Arthritis and Rheumatism* 18, 111-120.
- Goldenberg, D.L., R.W. Simms, A. Geiger, A.L. Komaroff. 1990.  
High frequency of fibromyalgia in patients with chronic fatigue seen in a primary care practice.  
*Arthritis and Rheumatism* 33(3), 381-387.
- Halperin, J.J. M. Golightly. 1992.  
Lyme borreliosis in Bell's palsy.  
*Neurology* 42, 1268-1270.
- Hanrahan, J.P., J.L. Benach, J.L. Coleman, E.M. Bosler, J.C. Grabau, D.L. Morse. 1984.  
Epidemiologic features of Lyme disease in New York.  
*The Yale Journal of Biology and Medicine* 57, 643-650.

- Hsu, V.M., S.J. Patella, L.H. Sigal. 1993.  
 "Chronic Lyme disease" as the incorrect diagnosis in patients with fibromyalgia.  
*Arthritis and Rheumatism* 36(11), 1493-1500.
- Huppertz, H.I. 1990.  
 Childhood Lyme borreliosis in Europe.  
*European Journal Pediatrics* 149, 814-821.
- Kahneman, D., Slovic, P., Tversky, A. 1982. *Judgement Under Uncertainty: Heuristics and Biases*, New York, N.Y. , Cambridge University press.
- Kaplan, R.F., M.E. Meadows, L.C. Vincent, E.L. Logigian, A.C. Steere. 1992.  
 Memory impairment and depression in patients with Lyme encephalopathy: comparison with fibromyalgia and nonpsychotically depressed patients.  
*Neurology* 42, 1263-1267.
- Magarelli, L.A., Ryan, R.W., Tilton, R.C., Hardin, J.A., Niejadlik, D.C., Sweeney, A.H., Cartter, M.L., Checko, P.J., Mshar, P.A., Chaski, H.C., Hadler, J.L., and Steere, A.C. 1988.  
 Lyme disease-Connecticut.  
*Journal of the American Medical Association* Vol. 259: 1147-1148.
- Magid, D., B. Schwartz, J. Craft, J.S. Schwartz. 1992.  
 Prevention of Lyme disease after tick bites: a cost effectiveness analysis.  
*The New England Journal of Medicine* 327(8), 534-540.
- Magnarelli, L.A., J.F. Anderson, C.S. Apperson, D. Fish, R.C. Johnson, W.A. Chappell. 1986.  
 Spirochetes in ticks and antibodies to *Borrelia burgdorferi* in white-tailed deer from Connecticut, New York state, and North Carolina.  
*Journal of Wildlife Diseases* 22(2), 178-188.
- Neapolitan R.E. 1990. *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*, John Wiley and Sons, Inc., New York.
- Petersen, L.R., A.H. Sweeney, P.J. Checko, L.A. Magnarelli, P.A. Mshar, R.A. Gunn, J.L. Hadler. 1989.  
 Epidemiological and clinical features of 1149 persons with Lyme disease identified by laboratory-based surveillance in Connecticut.  
*The Yale Journal of Biology and Medicine* 62, 253-262.

- Peisman, J., A. Spielman. 1979.  
Host-associations and seasonal abundance of immature *Ixodes dammini*  
in southeastern Massachusetts.  
*Annals of the Entomological Society of America* 72(6), 829-831.
- Piesman, J., T.N. Mather, R.J. Sinsky, A. Spielman. 1987.  
Duration of tick attachment and *Borrelia burgdorferi* transmission.  
*Journal of Clinical Microbiology* 25(3), 557-558.
- Piesman, J., T.N. Mather, G.J. Dammin, S.R. Telford III, C.C. Lastavica,  
A. Spielman. 1987.  
Seasonal variation of transmission risk of Lyme disease and human Babesiosis.  
*American Journal of Epidemiology* 126(6), 1187-1189.
- Piesman, J. 1993.  
Dynamics of *Borrelia burgdorferi* transmission by nymphal  
*Ixodes dammini* ticks.  
*Journal of Infectious Diseases* 167, 1082-1085.
- Rees, D.H.E., J.S. Axford. 1994. Evidence for Lyme disease in urban  
park workers: a potential new health hazard for city inhabitants.  
*British Journal of Rheumatology* 33, 123-128.
- Schulze, T.L., G.S. Bowen, M.F. Lakat, W.E. Parkin, J.K. Shisler. 1986  
Seasonal abundance and hosts of *Ixodes dammini* (Acari: Ixodidae)  
and other Ixodid ticks from an endemic Lyme disease focus in New Jersey,  
USA.  
*Journal of Medical Entomology* 23(1), 105-109.
- Shih, C-M., A. Spielman. 1993.  
Accelerated transmission of Lyme disease spirochetes by partially fed  
vector ticks.  
*Journal of Clinical Microbiology* 31, 2878-2881.
- Sigal, L.H. 1990.  
Summary of the first 100 patients seen at a Lyme disease referral center.  
*The American Journal of Medicine* 88, 577-581.
- Sigal, L.H., S.J. Patella. 1992.  
Lyme arthritis as the incorrect diagnosis in pediatric and adolescent  
fibromyalgia.  
*Pediatrics* 90(4), 523-528.

- Spiegelhalter, D.J., Dawid, A.P., Lauritzen, S.L., Cowell, R.G. 1993.  
Bayesian Analysis in Expert Systems.  
*Statistical Science* 8(3), 219-233.
- Stafford, K.C. III, L.A. Magnarelli. 1993.  
Spatial and temporal patterns of *Ixodes scapularis* (Acari: Ixodidae)  
in southeastern Connecticut.  
*Journal of Medical Entomology* 30(4), 762-771.
- Steere A.C., S.E. Malawista, J.A. Hardin, S. Ruddy, P.W. Askenase,  
W.A. Andiman. 1977.  
Erythema chronicum migrans and Lyme Arthritis.  
*Annals of Internal Medicine* 86(6), 685-698.
- Steere, A.C., A. Gibofsky, M.E. Patarroyo, R.J. Winchester, J.A. Hardin,  
S.E. Malawista. 1979.  
Chronic Lyme arthritis.  
*Annals of Internal Medicine* 90, 896-901.
- Steere, A.C., N.H. Bartenhagen, J.E. Craft, G.J. Hutchinson, J.H. Newman,  
D.W. Rahn, L.H. Sigal, P.N. Spieler, K.S. Stenn, S.E. Malawista. 1983.  
The early clinical manifestations of Lyme disease.  
*Annals of Internal Medicine* 99, 76-82.
- Steere, A.C., E. Taylor, M.L. Wilson, J.F. Levine, A. Spielman. 1986.  
Longitudinal assessment of the clinical and epidemiological features  
of Lyme disease in a defined population.  
*The Journal of Infectious Diseases* 154(2), 295-300.
- Steere, A.C., E. Taylor, G.L. McHugh, E.L. Logigian. 1993.  
The overdiagnosis of Lyme disease.  
*JAMA* 269(14), 1812-1816.
- Wahlberg, P., H. Granlund, D. Nyman, J. Panelius, I. Seppala. 1993.  
Late Lyme borreliosis: epidemiology, diagnosis and clinical features.  
*Annals of Medicine* 25, 349-352.
- Waylonis, G.W., W. Heck. 1992.  
Fibromyalgia syndrome.  
*American Journal of Physical Medicine and Rehabilitation* 71(6), 343-348.

Whelton, C.L., I. Salit, H. Moldofsky. 1992.  
Sleep, Epstein-Barr virus infection, musculoskeletal pain, and depressive  
symptoms in chronic fatigue syndrome.  
*The Journal of Rheumatology* 19(6), 939-943.