A Knowledge-Based Mastitis Diagnostic System for Dairy Participants in USA

Tae Woon Kim*·Jae Deuk Lee**

지식베이스에 의한 젖소 유방염 진단체계 개발

김태운*·이재득**

Abstract

The major economic health problem of dairy cattle is mastitis which can affect 10 to 50% of cow-quarters. This health problem is difficult for many dairy farmers and health advisors to understand, diagnose and control. Without special laboratory testing, most mastitis is overlooked. Estimates of annual mastitis cost per cow vary from $50 to $200. For the nearly 9 million cows in the United States, annual loss to the dairy industry amounts to over one billion.

A knowledge-based decision aid has been developed to evaluate mastitis data retrieved electronically from two of nine U.S. regional dairy records processing centers. Heuristic rules to diagnose herd mastitis problems were collected and incorporated into the system from various domain experts. This system imports database files for individual cows and queries the user for appropriate herd management information. It allows users to select mastitis control schemes with various degrees of aggressiveness and teaches commonly accepted mastitis control practices.

Keywords : Mastitis (유방염), Somatic Cell Count (체세포수), Fuzzy set, Database, Knowledge-based System

* 경성대학교 산업공학과
** 한성대학교 산업공학과

Ⅰ. Introduction

Mastitis, inflammation within the mammary gland in dairy cattle, is a disease that is difficult for dairy farmers and health advisors to understand, diagnose, and control. As the pathogens that cause mastitis are extremely diverse, treatments are specific to the cause. Farm advisors with the ability and inclination to understand mastitis records are rare. More uncommon are those that can organize diagnostic information from several sources, determine cause and effect relationships, and design herd mastitis control systems. Without good advice, farmers with herd mastitis problems often turn to antibiotic therapy as their choice of mastitis control. Therapy of this nature is of questionable value, is often costly, and raises the risk of condemnation of bulk milks with detectable antibiotic residues. Estimates of annual mastitis costs per cow vary from more than $50 to $200. Using $100 per cow year for the 9 million cows in the USA, the loss to the dairy industry amounts to over one billion.

Somatic cells in milk are a mixture of secretory epithelial cell and leukocytes (white cell). As the cell count increases dramatically, it generally represents an increase in polymorphonuclear leukocytes, which search out bacteria and their toxins in an effort to protect the udder from damages. Leukocytes are attracted to the cite of infection by a process called chemotaxix. The chemical stimuli can be messengers from injured cells, bacterial toxins themselves, or products of an inflammatory reaction. Somatic cell levels increase when a cow is infected, as a cow ages (because of greater probability of infection), and when a lactation is long, as a result of infection, production drops below 20 lb/day.

Based on control practices, organisms that cause mastitis can be divided into two categories namely environmental and contagious bacteria. The process of infection is as the following: The largest class of cows in a herd are uninfected; some cows are infected but go unnoticed, many of thesereturn to uninfected state by spontaneous cures; a minority of the cows in a herd become clinically infected that is observable infections without testing. Cows with environmental pathogens can progress through these three stages in hours to several days. Contagious infections often progress through these states in weeks to years with several months being typical. Most chronic mastitis infections are caused by contagious pathogens, which are often subclinical and go unobserved by farm workers. It is hard to predict mastitis infection before bacterial culture test. However, Somatic Cell Count (SCC), lactation number, and milk production in Dairy HerdImprovement Association (DHIA) reports are used to quantify and assess subclinical mastitis inexpensively once a month or less under most commercial conditions (Berning and Shook, 1992).

In recent years, powerful computer techniques have been used to address farm management challenges. Artificial intelligence is an example of one of these decision aids. Decision support systems or expert systems have been presented by Domecq et al. (1991), Schmisseur and Gamroth (1993), and

Pellerin et al. (1994) for dairy management advice. For mastitis management, Hogeveen et al. (1994), Allore et al. (1995), and Heald et al. (1994) have reported decision support systems. Mostly, decision support systems provide data analysis and classification of herd mastitis problems. Limited application of knowledge-based systems has been made to the prediction of mastitis in individual cows based on both DHIA and herd management data.

The purpose of this research was to develop the Mastitis Control System (MCS) as an effective tool for a mastitis control advisor. This tool should improve the management value of individual farm SCC value. Advisors can retrieve herd SCC data, evaluate SCC data, uniformly collect pertinent ancillary farm management data, predict weaknesses in individual farm mastitis control schemes, teach specific mastitis control procedures to correct unique herd problems, and decrease the reliance on antibiotic therapy. The objectives of this work were:

1). To develop a hypertext style database about mastitis to advise dairy farmers and farm consultants.

2). To construct an expert system that can assist veterinarians, extension agents, and farm consultants in the collection, simplification, analysis, and interpretation of DHIA data, laboratory test result, clinical disease, and symtomatic data in a consistent and methodical manner.

Ⅱ. Structure of Mastitis Control System

The MCS consists of several modules: hypertext style knowledge base, accessing mainframe database, and diagnosing mastitis problem. In order to access mainframe, the program has communication software unique to each Dairy Records Processing Center (DRPC), C++ program to preprocess ASCII data file retrieved from the DRPC, and an interface program written with Toolbook(Asymetrix Corp., 1996). The program only requires the use of a mouse or similar pointing device. All data reduction, analyses, heuristic evaluation (including fuzzy logic), graphing, modem communications, and queries are accomplished within the customized Toolbook application. The system is composed of three main modules shown in Figure 1: downloading DHIA mainframe databases, evaluation of the mastitis problem, and retrieval of existing knowledge bases.

2.1 Downloading DHIA Databases

DHIA, farmer-owned cooperatives is a regional dairy records processing center. Milk components analyses are conducted monthly on each cow, the data is stored in a regional dairy records processing center, and paper reports are sent to each farm. Monthlyrecords may include the date of test, daily milk ield, fat and protein percentage, and SCC. Factors affecting the records are also recorded and stored. Through a modem, data items from DHIA mainframe can be downloaded automatically using script language in the Toolbook software. ASCII data file from DHIA were preprocessed before analysis, organization, and summarization in the MCS. This module is one step process, entirely accomplished by the action of pressing one button.

2.2 Evaluation of the Mastitis Problem

The mastitis control system considers data from DHIA reports, generalized clinical mastitis symptoms, and general herd management information. Lactation number, milk production, and SCC in DHIA reports are the variables used to quantify and assess subclinical mastitis. As the value of SCC ranges from hundred-thousands to several millions, a logarithmic phenomena that is transformed to the log base 2 is used for many analyses. Transformed data termed Somatic Cell Score(SCS) in contrast to untransformed data termed SCC. The SCS data is retrieved from the regional DRPC; the clinical and management information is entered in the MCS diagnostic survey from farmer input using interface screen.

The evaluation module analyzes individual cow and general herd data, and graphically depicts results of analyses. Other graphs and tables show the trend of SCS according to lactation and test date and provide the name and related data of cows ranked by the highest SCS in the current and previous months. Heuristic rules to diagnose herd mastitis problems have been collected and incorporated into the system from various domain experts. After diagnosis, users are also able to select mastitis control schemes. These control schemes are commonly accepted mastitis control practices developed by domain experts.

The developed system is intended to help dairy farmers and dairy advisors who might be computer novices and do not have enough time to work on the PC. Thus, it is composed of command buttons which respond to the mouse-click from a user. It is primarily designed and implemented on the ideas of user-friendly and ease of use.

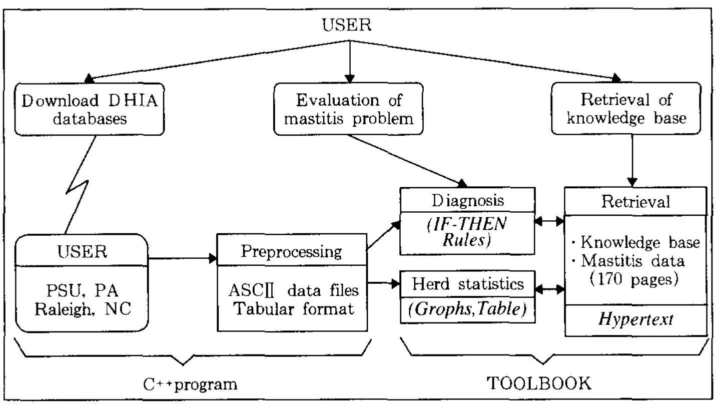


Figure 1. Structure of Mastitis Control System.

2.3 Retrieval of Hypertext-style Knowledge Bases

About 170 pages of screens are inter-linked to provide ad hoc information retrieval on specific mastitis topics selected by the user. Screens contain text, pictures of laboratory cultures, illustrations, and graphical data to enhance users' understanding. This module is designed to be user-friendly and requires minimal computer skills. Most program functions are obtained by moving a pointing device to position the pointer on on-screen buttons. Figure 2 displays an example of hypertext-style screen showing the relationship between SCC and SCS.

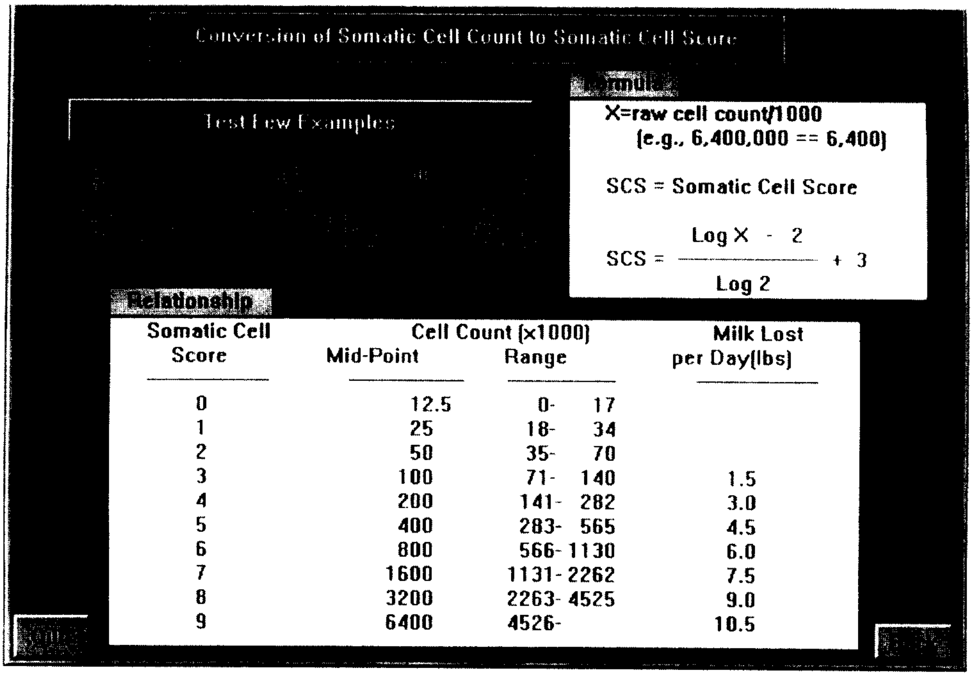


Figure 2. A Hyper-text Style Screen. (It shows relationship between SCC and SCS.)

Ⅲ. Knowledge Representation and Diagnostic Rule Inference

3.1 Knowledge Acquisition

Knowledge acquisition is the most crucial aspect in building an expert system. It is the process of extracting, structuring, and organizing domain knowledge from available sources. This is a painstaking task requiring a knowledge engineer familiar with the structure and mechanism of the system to interview a domain expert in mastitis. The goal of the knowledge engineer is to elicit information that will lead to better performance. In order to ask questions that will elicit this sort of information, the knowledge engineer should have a degree of mastery over the domain and a sense of the kind of knowledge. Knowledge required for the evaluation of mastitis can be classified as facts from the domain and transformed information from the facts, in which the former corresponds to declarative knowledge and the latter corresponds to procedural or operational knowledge.

Declarative knowledge about mastitis is both general and specific. The general knowledge is acquired from national mastitis literature databases, journals, and interviews with veterinarians, dairy consultants, farmers, and other professionals. Specific knowledge about a particular herd is acquired by downloading databases from DRPCs using a modem. Procedural knowledge is transformed or interpreted information from existing declarative knowledge to be used as a decision aid. It is acquired from evaluation results in the MCS. Declarative knowledge, which is stored as a separate ASCII data file, can be uploaded in the MCS to predict mastitis for a specific farm or cows. On the contrary, procedural knowledge is programmed within the MCS and access declarative knowledge to diagnose mastitis. The following section describes how to deal with the issue of uncertainty, a major obstacle for the generation of procedural knowledge, and inference mechanism.

3.2 Uncertain Knowledge Representation using Fuzzy Logic

Much of the information related to mastitis is asymptotic, and not clearly quantified, but is uncertain or fuzzy. Thus, conventional knowledge representation techniques based on predicate logic and related methods are not suitable for representing inexact and qualitative knowledge. In predicate logic, a proposition is either true or false; no gradation of truth or membership are allowed. The fuzzy logic approach for the representation of uncertainty is based on the fuzzy membership function, which specifies the degree to which any output satisfies the conditions (Zadeh, 1983).

Organisms that cause mastitis are loosely divided into two primary groups termed contagious bacteria (*Strep. ag.* and *Staph. aureus*) and environmental bacteria (*Strep.* non-ag, Coliforms, and Klebsiella, to name a few). Producer responses help to determine whether or not contagious mastitis may be present in a herd. Figure 3 show the user input screen for a contagious case.

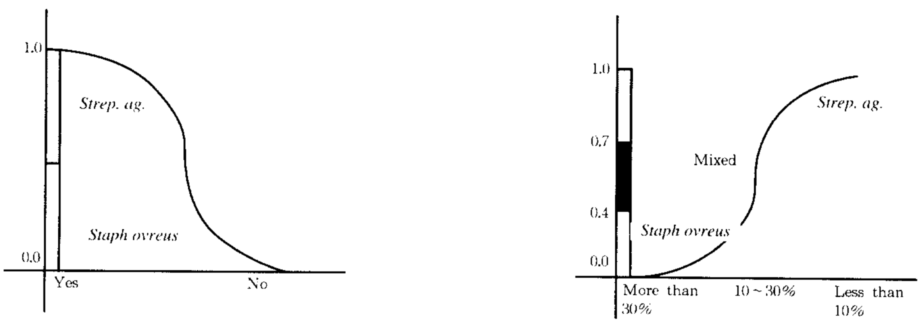
Rules for Contagious(strep ag/Staph aureus)

* Click on Yes or No box(☒)

	Yes	No	
1. Did milk from clinical cases cultured over the last year find Strep ag in more than 1 cow?	☐	☒	
2. Are all cows teat dipped at every milking w/an effective product?	☒	☐	3 more years
3. Are all cows given dry cow therapy when turned dry?	☒	☐	3 more years
4. Were there any cows brought into the herd?	☐	☒	
5. Were cows with clinecal mastitis tested for bacterial type?	☒	☐	more than 30%
6. Is a common sponge or wash cloth used to clean udders before milking?	☐	☒	
7. Of the cows with persistent clinical mastitis, are there any with two or three quarters infected?	☐	☒	
8. Do cows with persistant mastitis generally respond to antibiotic therapy and not return with mastitis in one month?	☒	☐	
9. Co cows with mastitis tend to exceed SCS 7?	☐	☒	

Figure 3. User Input Screen regarding Rules for Contagious Mastitis.

The rules related to the uncertain status consist of some conditions on the range of acceptable outputs. Every condition on the range of possible output is associated with membership function, which takes a value between 0 and 1. Membership function can be established through an interview of domain experts to mastitis problems. For example, the membership function of rule 1 in Figure 3 is drawn from the expert's input: When the expert's response to rule 1 is "yes," it means "most likely *Strep. ag.*"; when the response to rule 1 is "no," it means "most likely *Staph. aureus.*" Every membership function corresponds to a rule(s) and is generated in a similar manner. The membership function corresponding to rule 1 and rule 5 is illustrated in Figure 4.



Rule 1.(Did milk from clinical cases cultured over the last year find *strep. ag.* in more than I cow?)

Rule 5.(Were cows with clinical mastitis tested for bacterial type?)

Figure 4. Examples of Membership Functions for Contagious Mastitis.
(Rule 1 and Rule 5 corresponds to the cases in Figure 3.)

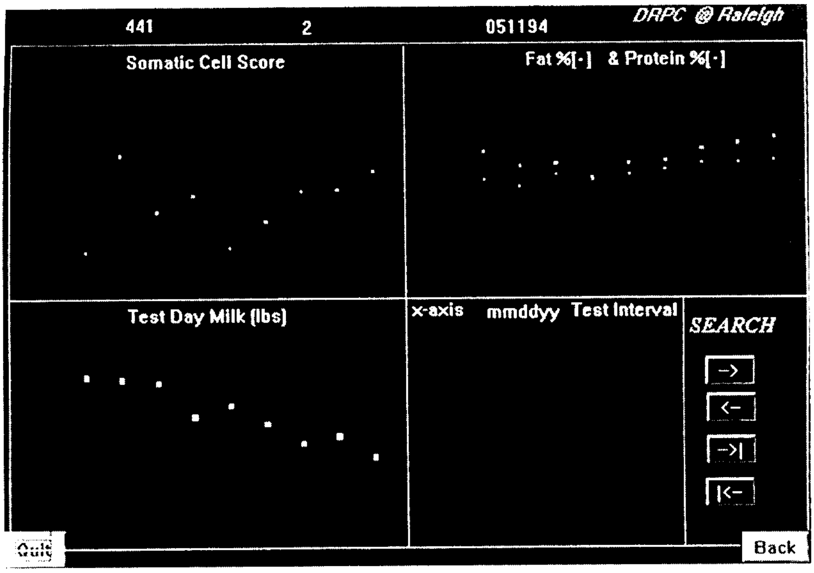


Figure 5. Graphical Data Display for Individual Cow from DHIA.
(The number 1 to 9 in x-axis correspond to the test date.)

In order to evaluate contagious mastitis problem, the group of rules in Figure 3 should be jointly satisfied. In this case, th proposed output (*Strep. ag.* or *Staph. aureus*) is evaluated for each rule, and then assign the performance score equal to the lowest one among the individual rules. Individual performance score corresponding to conditional rule is defined. The expression, "IF the inputs satisfy condition C, then the output must conform to the constraint S" is represented by:

Individual performance score (p) = max [$1 - a_c$ (input), a_s (putput)]

where a_c, a_s are membership functions of input and output variables respectively.

The overall performance score for a conjunctive system of rules can function exactly as if it was obtainedd from a single rule.

Sometimes, there exists nested relationships among the rules. The nested rules are logically the same as simple rules, that is "if a, then b then c" is the same as "if a and b, then c." The nested relationship is found between rule 1 and rule 5 in Figure 3. The combined meaning of the membership function from two rules is:

IF rule 1 is *false (no)*

and

IF rule 5 is *true (yes)*

and

more than 30%

THEN it is *most likely not strep ag.*

3.3 Inference of Diagnostic Rule

Knowledge for diagnosing mastitis has been summarized as IF-THEN rule structure from domain expert's experience. Parameters used for the rules are mostly DHIA data items, and a few of them are from user input screen shown in Figure 3. The inference method is forward chaining which starts from DHIA data items and evaluate them based on the rule structure. Inference systemcombines both facts and knowledge base. After a rule which satisfies the conditional part from the given fact is executed, a new fact in the conclusion part of the rule is stored in the fact base. This process continues until diagnosis of mastitis is completed for a cow or group of cows. The knowledge base is composed of around 200 diagnostic rules. Each rule can be executed using the facts from farm DHIA data and user input. The system is designed to maximize the usage of downloaded data and to minimize user input for easy use by dairy farmers. A typical forward chaining repeats the sequence of pattern matching, conflict resolution, and firing. For the pattern matching, rules and DHIA database are evaluated during execution, then the matched data and rules are stored in the working memory in the Toolbook. Conflict resolution is performed according to the predefined priority among decision variables and a set of rules.

Rule representation and inference have been programmed using Toolbook as a software vehicle. Toolbook is a script language especially designed for a multimedia, a training program, and an educational development tool. It has hot-key function which can be inter-connected to other pages or keywords similar to the function of HTML language. Also, image and graphical functions can be easily supported within Toolbook package. As the main purpose of MCS is a decision aid for dairy participants who are computer novices, Toolbook makes it possible to combine images, graphs, tables, hyper-text, user input screen, and inference mechanism without any other software support.

Ⅳ. Implementation

Basically, MCS can support three functions. First, it contains and provides common knowledge about mastitis and appropriate control programs. Second, it shows herd-specific graphs and tabular data which has been downloaded from DHIA databases. Figure 5 shows graphical data of SCC, milk production, percentage of fat or protein, and test date during 9 month period for individual cow. Third, it can evaluate the mastitis problem from a whole herd or individual cow perspective. The system utilizes downloaded DHIA data and user input relative to the herd status. Also, the system can evaluate mastitis status both whole herd and individual cow perspective.

To evaluate a whole herd, the system generates a frequency table composed of first infected and severely infected cows from DHIA data. By applying domain expert's knowledge, the herd mastitis status is evaluated as contagious, environmental, and/or mixed. The system can also suggest which lactation group or groups have significant problems, i.e., lactation 1, 2, and/or 3 or plus, and early, mid, or late lactation (days in milk).

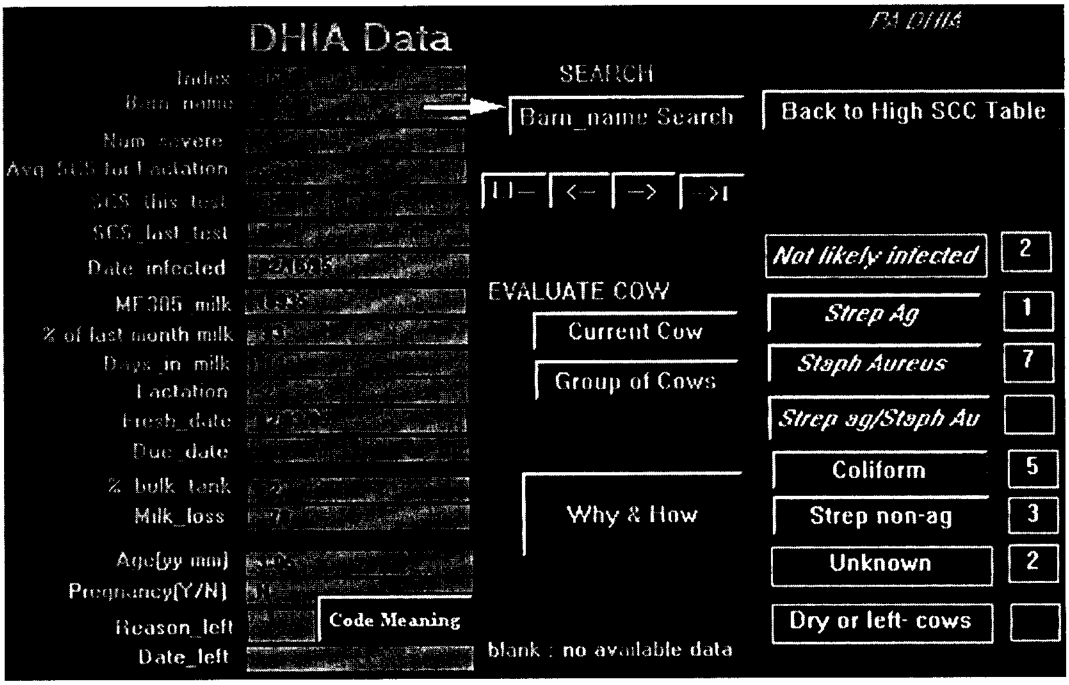


Figure 6. A Screen for Diagnosing Mastitis
(Number on each category on the right-hand side represents
the summary of diagnostics results after evaluation of 20 cows.)

In the individual cow evaluation mode, the user can evaluate as many cows as desired by specifying the number of cows on the user input screen. Eight diagnostic categories are predefined as shown in Figure 6. The categories are: not likely infected, *Strep Ag.*, *Staph Aureus*, mixed (*Strep. Ag.* and *Staph Aureus*), Coliform, *Strep.* non-ag, unknown, not tested (dry or left cows). The number next to each category shows the diagnostic results of 20 cows. An explanation function is provided by pressing the "Why & How" button to determine how each cow's mastitis status was diagnosed from the available data. Figure 7 shows the summary data related to the diagnostic results corresponding to the 20 cows.

The potential of the program attracted the attention of national dairy science professionals, mastitis advisors, and veterinary practitioners when the prototype MCS was presented (Heald and Kim, 1994).

Herd ID : 23140185 — History of Evaluation (Group of Cows)

No	Cow Name	Cow Index	Lact	DIM	Fresh Date	Recent SCS	Last SCS	Diagnostic Result
1	415	1537	2	50	12/29/94	6.18	8.67	Coliform
2	752	1305	6	89	11/20/94	8.13	8.30	Coliform
3	900	1411	4	46	01/02/95	8.76	1.11	Coliform
4	996	1518	2	236	06/26/94	7.99	7.60	Staph aureus
5	810	1371	5	16	01/04/95	8.21	4.30	Strep non-ag
6	740	1293	6	70	12/09/94	6.64	7.80	Coliform
7	641	1214	7	77	12/02/94	6.09	6.46	Coliform
8	494	1557	1	367	02/15/94	4.81	5.87	Staph aureus
9	410	1519	2	130	10/10/94	6.60	5.61	Staph aureus
10	527	1618	1	39	01/09/95	1.71	7.25	unknown
11	504	1608	1	150	09/20/94	2.57	5.80	Not Likely Infected
12	303	1443	3	156	09/14/94	3.06	2.77	Not Likely Infected
13	785	1332	5	149	09/21/94	6.19	4.53	Staph aureus
14	748	1301	6	230	07/02/94	6.72	5.68	Staph aureus
15	864	1389	4	152	09/18/94	7.10	7.09	strepag
16	805	1349	5	14	02/03/95	5.96	0.00	Strep non-ag
17	521	1614	1	69	12/10/94	1.44	4.97	unknown
18	424	1539	1	397	01/16/94	5.27	5.63	Staph aureus
19	920	1450	3	281	05/12/94	3.55	4.03	Staph aureus
20	446	1548	2	14	02/03/95	5.71	0.00	Strep non-ag

Figure 7. Diagnostic Summary Results after Execution of 20 Cows.

V. Conclusions

An expert system is now available to help DHIA dairy farmers(specifically with SCC records) evaluate their mastitis control programs and learn how to improve their mastitis status. This program can be used with remote data files from herds that are members of the DRPCs at University Park, PA and Raleigh, NC. Novices in mastitis control can use the program to easily retrieve DHIA records and automatically evaluate these records with a significant savings of time and effort.

The result of having a computerized, single-site mastitis knowledge base should aid in the reduction of financial losses due to mastitis and improve the quality of milk marked by dairy farmers. Although the MCS was developed for herds belonging to the dairy record processing centers in Pennsylvania and North Carolina, it is essentially a generic system with global application. Once further verified under Pennsylvania conditions, the MCS could be used in other DRPCs throughout the United States upon modification of the interface program for retrieval from other databases.

References

Allore, H. G., L. R. Jones, W. G. Merrill, and P. A. Oltenacu. "A Decision Support System for Evaluating Mastitis Information," *J. Dairy Science*, Vol. 78 (1995), 1382-1398.

Asymetrix Corp., Toolbokk II, Asymetrix Corp., 1996.

Berning, L. M. and G.E. Shook, "Prediction of Mastitis using Milk Somatic Cell Count, N-Acetyl-b-D-Glucosaminidase, and lactose," *J. of Dairy Science*, Vol. 75 (1992), 1840-1848.

Domecq, J.J., R.L. Nebel, M. L. McGilliard, and A. T. Pasquino. "Expert system for evaluation of reproductive performance and management," *J. Dairy Science*, 74 (1991), 3446-3452.

Heald, C.W., and T. Kim, "A Computer Decision Aid to Help Identify Mastitis Causes and Teach Mastitis Control," Proceedings 33rd Annual Meetings of National Mastitis Council, Orlando, FL, January 31-February 2 (1994), 339-340.

Hogeveen, H., M.A. Varner, D.S. Bree, D.E. Dill, E.N. Noordhuizen-Stassen, and A. Brand, "Knowledge Representation Methods for Dairy Decision Support Systems," *J. Dairy Science* Vol. 77 (1994), 3704-3715.

Pellerin, D., R. Levallois, G. St-Laurent, and J. P. Perrier, "LAIT-XPERT VACHES: An Expert System for Dairy herd Management," *J. Dairy Science*, Vol. 77 (1994), 2308-2317.

Schmisseur, E., and M.J. Gamroth, "DXMAS: An Expert System Program Providing Management Advice to Dairy Operators," *J. Dairy Science*, Vol. 76 (1993), 2039-2049.

Zadeh, Lofti A, "Common Sense Knowledge Representation based on Fuzzy Logic", *Computer*, Vol. 17, No. 10 (1983), 61-65.