2003

Collaborative interpreting for knowledge discovery.

Daisy Yin-Shang Wong
University of Alabama at Birmingham

COLLABORATIVE INTERPRETING FOR KNOWLEDGE DISCOVERY

by

DAISY Y. WONG

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2003

UMI Number: 3101537

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT OF DISSERTATION
GRADUATE SCHOOL, UNIVERSITY OF ALABAMA AT BIRMINGHAM

Degree Ph.D. Program Computer and Information Sciences

Name of Candidate Daisy Y. Wong

Committee Chair Warren T. Jones

Title Collaborative Interpreting for Knowledge Discovery

The purpose of Knowledge Discovery and Data-mining (KDD) systems is to uncover novel, non-trivial patterns; however, these systems also discover many useless patterns. The interpretation and evaluation of patterns requires domain expertise. The process is inherently subjective because experts differ in experience and knowledge. In most KDD research, interpretation/evaluation has been done by the researchers or by a single expert. In this research, we introduce Collaborative Interpreting (CI), which uses the collective expertise of a diverse group of experts to evaluate data-mining results. Our hypothesis is that the collaborative interpretation of the group would be better than that of any single expert[1].

This research involves the development of a World Wide Web (WWW)-based Collaborative Interpreting System (CIS), which employs an iterative group decision-making process based on the Delphi Method. The CIS compiled the interpretations of a panel of geographically and temporally distributed experts on a continuous basis. The application domain used was hospital infection surveillance. The panel consisted of two infectious disease physicians, two critical care physicians, a pharmacist, and a clinical microbiologist. Patterns were generated for five months by the Data Mining Surveillance

---

[1] Turoff M: Computer-Mediated Communication Requirements for Group Support. Journal of Organizational Computing 1991; 1: 85-113.

System (DMSS)[2] using retrospective clinical laboratory infection-control data from the University of Alabama at Birmingham Hospital. For each month, the panel used CIS to judge the importance of the patterns independently and then collectively with regard to epidemiology investigation priorities.

Group and individual judgments were compared to those of a criterion judge who is an infectious disease physician with more than thirty years of infection control experience. In addition to the patterns, he was provided with patient chart review data. The results showed that the panel's collective judgments were generally in greater agreement with the criterion than any individual member of the panel. The results confirmed the hypothesis and showed that interpretation is a more complex problem than generally assumed.

The findings are encouraging that CIS can 1) utilize collective expertise to enhance the interpretation of KDD results, 2) enrich the knowledge base of experts through collaboration with other experts, and 3) potentially help refine data-mining algorithms to reduce generation of trivial patterns.

---

[2] DMSS was originally developed at the University of Alabama at Birmingham by S. Brossette.

DEDICATION

To my family,

and to all children who do not have a family.

May God bless them with one soon!

ACKNOWLEDGMENTS

First of all, I would like to thank my two mentors, Dr. Warren Jones and Dr. Stephen Moser, for their continuous support, encouragement, and patience through this educational journey. I learned about the field of Knowledge Discovery and Data Mining in Dr. Jones' Data Mining class. From that class, I realized that the interpretation of data mining results is not a simple problem. Dr. Jones always keeps an open mind to different ideas. When he was helping me to formulate my research focus, he kept in mind my professional experience in computer human interface and software engineering, and my interest in the field of Computer Supported Collaborative Works. He is always available to help me to overcome obstacles. He is also a proponent of collaboration with other departments. He was instrumental in introducing me to Dr. Moser of the Department of Pathology, who kindly agreed to be my co-advisor. This research would not have been possible without Dr. Moser's support and guidance. He helped me to understand the domain of Infection Surveillance and to navigate the complex organizational structure of the UAB medical center. He helped to recruit experts as subjects for the experiment. He also provided the needed resources to conduct the research and served as the moderator during the experiment. He has been very generous with his time in going through all of the details of the research with me. I am very grateful to both of them.

I would like to thank Dr. J. Michael Hardin for pointing me to the Delphi Method and for providing me guidance in the statistical analysis. I also thank him for serving on my committee.

Dr. Barrett Bryant is always willing to listen to my rants about database and J2EE. He helped me to make contact with the appropriate Sun representative to gain a better understanding about Enterprise Java Beans (EJB) in order to improve my prototype's performance. Dr. Robert Hyatt has accommodated my computing needs during my search for the appropriate platform to develope the CIS prototype. He and Mr. Bruce Williams have gone out of their way to help me, including setting aside a machine for me to use, which gave me as much freedom as possible within the confines of the department firewalls. I thank both Dr. Bryant and Dr. Hyatt for taking the time to serve on my committee.

I want to express a sincere "thank you" to Dr. Stephen Brossette, who invented the special Data Mining Surveillance System (DMSS) for Infection Surveillance. His research with Dr. Jones and Dr. Moser provided the environment and the data-mining engine for my work. He generated the patterns for the experiment. He helped me to understand the DMSS patterns and the medical language, and he provided me with advice during the experiment.

I would also like to thank Dr. Kirk Avent for taking the time to serve as the criterion judge. He was generous with his time and advice, helping me with the user interface design and the patient chart review.

Dr. Alan Stamm, Chairman of Infection Control Committee, was very kind in helping me to understand infection control in the hospital. He invited me to attend his committee meetings as an observer and to interview his Infection Control Practitioners. He also allowed me to interview him to understand how an infection control specialist might interpret the DMSS patterns. He even allowed me to audiotape the interview for

detailed review later. He has been a valued resource in my research, and I want to thank him for his advice and time.

I would also like to thank Dr. Murray Turoff, the "Father of Delphi," for taking the time to answer my e-mail inquiry on the Delphi method and for sharing references of Delphi studies with me.

I want to express my appreciation to the panel of experts who volunteered their time and expertise. I regret that I cannot name them here because of the confidentiality agreement of the research protocol.

Thanks also to the team of chart reviewers, Vicki Fitzgerald, Nancy Olvey, April Calloway, and Stacy Aaron. The patient chart review task was indeed a very humbling experience to me.

There are many others with whom I have discussed ideas throughout this research, and I thank them for listening to me. Their questions and comments helped me to think more clearly.

Finally, I want to thank my family for their love and support through this whole endeavor. Special thanks to my husband, Dr. Stephen L. Chew, for providing me with intellectual and emotional support at all times. He is my Behavioral Science and Research Methods consultant. He is also my editor-in-chief. He is always my source of inspiration and my sounding board for ideas.

I have tried to incorporate comments from everyone who has read this manuscript and made suggestions. Any mistakes in this document are my own fault. Thank you for reading this dissertation. This research was supported in part by funding from the Alabama Academy of Science and from the Aetna Foundation, Inc.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

CI	Collaborative Interpreting
CIS	Collaborative Interpreting System
CMC	Computer-mediated Communication
DBMS	Database Management System
DHTML	Dynamic Hypertext Markup Language
DMSS	Data Mining Surveillance System
EJB	Enterprise JavaBean
ICC	Infection Control Committee
ICP	Infection Control Practitioner
J2EE	Java 2 Enterprise Edition
JSP	Java Server Page
KDD	Knowledge Discovery in Databases
LIS	Laboratory Information System
RDBMS	Relational Database Management System
UABH	University of Alabama at Birmingham Hospital
WWW	World Wide Web

CHAPTER 1

INTRODUCTION

Knowledge Discovery in Databases and Data Mining (KDD), is defined as the extraction of non-trivial, previously unknown, and potentially useful information from a large volume of data [3], [22], [47]. It is an iterative and interactive process generally comprised of the following stages: 1) data selection, 2) preprocessing, 3) transformation, 4) data mining, and 5) output interpretation and evaluation [22], [3]. Figure 1 is a KDD process diagram showing the five stages, the input and output for each stage, and the iterative nature of the process. The figure is adapted from the KDD process diagram in [22].

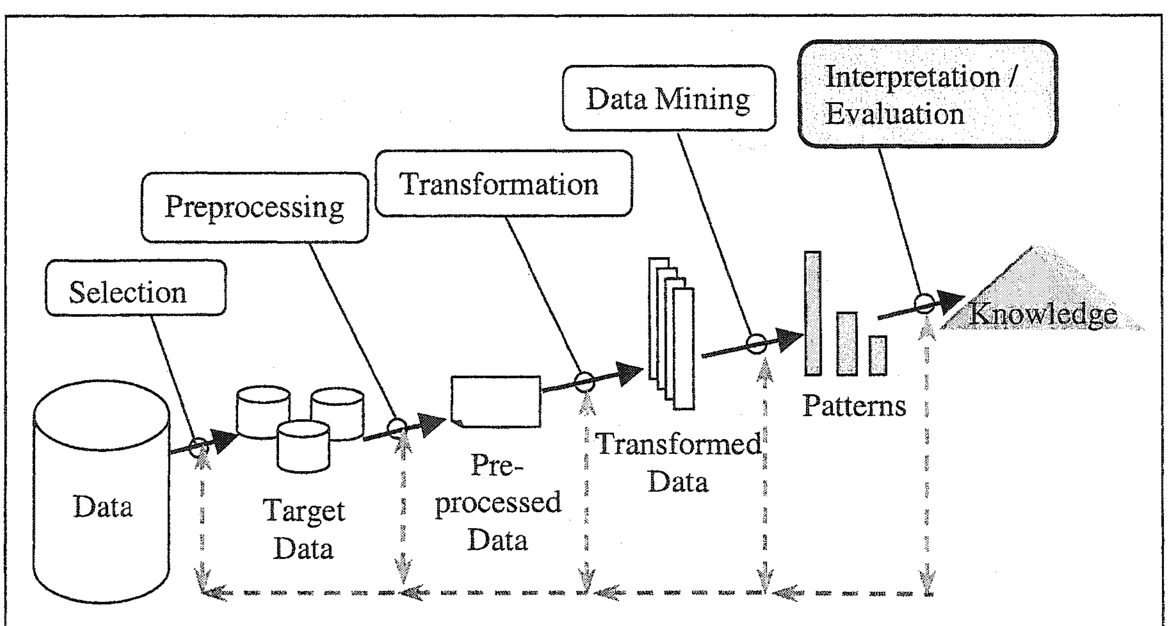


Figure 1. The KDD Process (Adapted from [22])

Stages 1 to 3 extract relevant data from organizational databases and prepare it for use in Stages 4 and 5. Stage 4, data mining, uses sophisticated computing algorithms, statistical analyses, and modeling techniques to uncover patterns and relationships hidden in the data. Stage 5, interpretation and evaluation, requires users to make sense of the discovered patterns.

The growing interest in the potential of KDD applications to unearth useful information from the rapidly growing amount of data collected at institutions and businesses has caused KDD research and commercial activities to flourish in the past ten years. For Stages 1 to 3, commercial database management systems developers, such as Oracle and Information Builder, are leading the efforts to address the issues at these stages with the development of data warehouse[3] and datamart[4] technologies which integrate KDD with the traditional database products. For the rest of the KDD research activities, the vast majority of the effort has focused on data mining techniques. This is evidenced by surveying the past eight International Conferences on Knowledge Discovery and Data Mining (KDD-95 to KDD-2002) [4, 10, 20, 21, 27, 28, 49, 58] and other KDD literature [3, 23, 44, 48]. In contrast, scant research has focused on the evaluation and interpretation of the data mining output to support decision-making. There is only a small amount of literature addressing this topic. This stage is highlighted with gray in Figure 1 because it is the focus of this dissertation research.

---

[3] A data warehouse is a large, integrated database extracted from operational databases for the purpose of decision support in businesses. The content of the warehouse is subject oriented, non-volatile, and time dependant [3].

[4] Datamarts are smaller, local data warehouses [3].

The remainder of this chapter delineates the motivation for this research. It defines Collaborative Interpreting (CI) and describes the experiment hypothesis, gives an overview of the Collaborative Interpreting System (CIS) developed for the research, and outlines the experiment used to examine the hypothesis. The final section of this chapter will provide an overview of the remaining chapters of this dissertation.

1.1 Motivation

The purpose of interpretation and evaluation is for end users to make sense of the discovered patterns and to determine what the patterns mean to their businesses or organizations. Research on interpretation has mostly focused on visualization of the extracted models or data, and on tools to support an end user in manipulating the mining parameters and interacting with the output data with various visualization tools. Data visualization is very useful for obtaining a global view of a resulting data set and highlighting important phenomena [59]; however, presenting the data in a visually salient way is not the same as interpretation by domain experts. The value of the discovered knowledge lies in its appropriate use, and this should be emphasized during interpretation and evaluation [66]. However, the use of the uncovered knowledge for decision-making is not a trivial process. This is because many of the current data mining tools report large amounts of useless or invalid data patterns besides the small amount of valuable knowledge. Large volumes of useless or trivial data can mask the useful knowledge, whereas invalid data masquerading as knowledge can lead to costly and disastrous decisions. In order to use the data mining results to support decision-making, interpretation requires sufficient domain expertise to perceive the nature of the unearthed relationships and the real implications of the uncovered patterns [2, 62, 67, 69].

Interpretation is inherently subjective because experts differ in experience and knowledge. If the discovered knowledge is really novel and previously unknown, it is highly unlikely that a single individual will possess all of the expertise needed to determine the validity and usefulness of the mined data for decision making. Organizations that want to gain a competitive edge using KDD tools realize that a single expert is often not sufficient to interpret the novel patterns discovered. Instead, a diverse team of data-mining-knowledgeable domain experts are needed to check and validate the mined results before the knowledge can be used to make decisions [62]. However, in most KDD research, interpretation and evaluation has been done by the researchers themselves or by a single expert [2, 19, 34, 67, 69, 4, 10, 20, 21, 27, 28, 49, 58]. Personal communications of the author with a number of data mining researchers indicated that researchers agree that using multiple domain experts to interpret the output is desirable. However, the use of multiple experts was not done in practice because it is difficult to get a number of experts together routinely to review their experimental output, and it is a laborious task to collect, aggregate, and organize the opinions of multiple experts, especially if it is done on a continuous basis. The ideal is to have a team of data-mining-knowledgeable domain experts to perform the interpretation. In reality, retaining a team of data mining knowledgeable domain experts is neither feasible nor practical for most organizations. Besides, people with such qualities are few and far between [62].

Uthurusamy [66] suggested that KDD should be a collaborative discovery process by a widely distributed team of experts and analysts. With the rapid growth of the

Internet and wide access to the World Wide Web, Uthurusamy's idea becomes feasible. Collaboration is especially critical for the interpretation and evaluation step. This idea forms the rationale for the CIS concept of this research. Computer-mediated communication pioneer, Turoff, suggested that "the results of collaboration are better than the result that could have been obtained by any single member of the group [63]." His conjecture provided the basis of this research's hypothesis.

Brossette [14] developed a special Data Mining Surveillance System (DMSS) at the University of Alabama at Birmingham to automate the extensive analysis of hospital clinical laboratory data in order to detect new trends of infection and drug-resistant bacteria. In an unpublished retrospective study of DMSS by Moser, one of the investigators, using the University of Alabama at Birmingham Hospital (UABH) data from April 1998 to November 1999, 40 out of 476 patterns were scored as interesting and worthy of investigation by a multi-disciplinary group from the hospital. Then two independent experts were asked to critically evaluate the patterns. They agreed with many of the 40 patterns originally selected, but disagreed over their importance for investigation and the exclusion of patterns that were not selected. This further illustrates that collaboration by a team of experts during interpretation would lead to a more meaningful screening of the findings and would reduce potential misses of important ones. This use of collaboration concurs with Farrington's view that a successful computer-based detection system requires "co-operation between epidemiologists, microbiologists, computer scientists and statisticians. Such an inter-disciplinary, collaborative approach can enhance existing methods, and improve the data collected in the process [19]. "

Experts in critical care medicine, surgery, infectious diseases, pharmacy, and clinical microbiology each have unique insights, and their knowledge and experience overlap with each other by varying magnitudes. Retaining a team of experts just for data interpretation, however, is neither feasible nor practical for hospitals. In addition, the experts are distributed both temporally and geographically. Therefore, an operational environment to compile and quantify the interpretation of DMSS patterns by multiple domain experts on an ongoing basis is needed. Such need provided an ideal domain environment to prototype the CIS of this research.

1.2 CI and Research Hypothesis

In this research, we introduce CI, which is a computer-mediated process that uses the collective expertise of a diverse group of geographically and temporally distributed experts to evaluate and prioritize data-mining results on a routine basis. The process facilitates the collection, compilation, quantification, and prioritization of the group's opinions on the discovered patterns. Our hypothesis was that collaborative interpretation by a group of diverse domain experts would be better than that by any single expert. We used Infection Control Surveillance at UABH as the application domain and used DMSS as the data-mining engine. We predict that collaborative interpretations would consistently be better than any individual's interpretations.

1.3 A Conceptual System for Collaborative Interpreting

The CIS is a web-based client-server application that provides an asynchronous interactive environment for collaborative interpreting as depicted in Figure 2. In this research, the actual implementation of the Data-mining Server and the Group Support Server ran on the same machine.

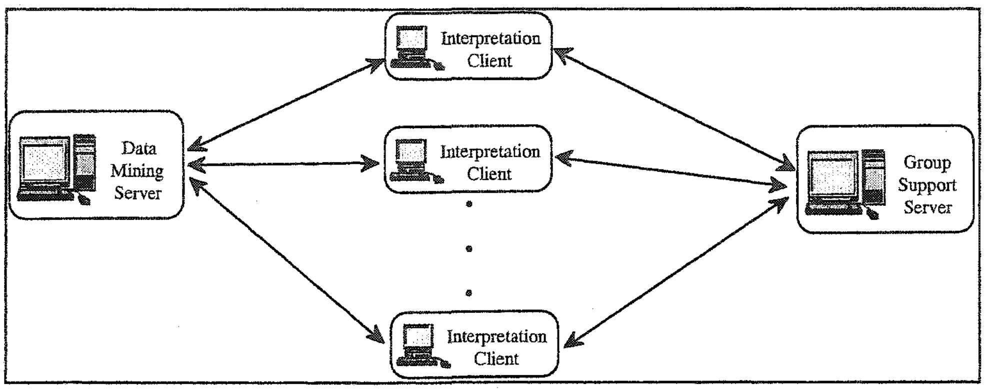


Figure 2. A CIS client-server diagram

It is conceptually divided into three subsystems according to functionalities:

1. The Client implements the user interface to present the patterns and the group's interpretations, and to collect the panelists' input to send to the servers.

2. The Group Support Server implements the group process logic, and stores and retrieves the group's interpretation in a relational database.

3. The Data-mining Server retrieves the patterns from the pattern database to present to the panelists, and supports drill down requests. "Drill down" means the ability to retrieve the raw data that are used to generate the data-mining patterns.

Figure 3 shows the conceptual software architecture of the CIS with an interface to a data-mining system. The actual design and implementation of the CIS prototype is described in Chapter 4.

1.4 Experiment Overview

The purpose of this experiment was to examine the hypothesis that the results of collaboration are better than the results that could have been obtained by any single member of the group alone. The experiment was conducted using the CIS prototype with patterns generated by DMSS using retrospective data from the UAB Hospital Clinical Laboratory Information System. The panel consisted of two infectious disease physicians, two critical care physicians, a pharmacist, and a clinical microbiologist. In addition, an expert who is also very knowledgeable in the data-mining technique used in this research participated in the experiment as an "out-group" panelist. The out-group panelist performed the same tasks as the panelists in the group, and saw the opinion of all group members as feedback. However, the in-group panelists did not see the opinion of the out-group panelist, and the group interpretation computation did not include input from the out-group panelist. The panelists did not know whether they were in the in-group or out-group. Group and individual judgments were compared to those of a criterion judge who is an infectious disease physician with more than thirty years of infection control experience. In addition to the patterns, he was provided with patient chart review data.

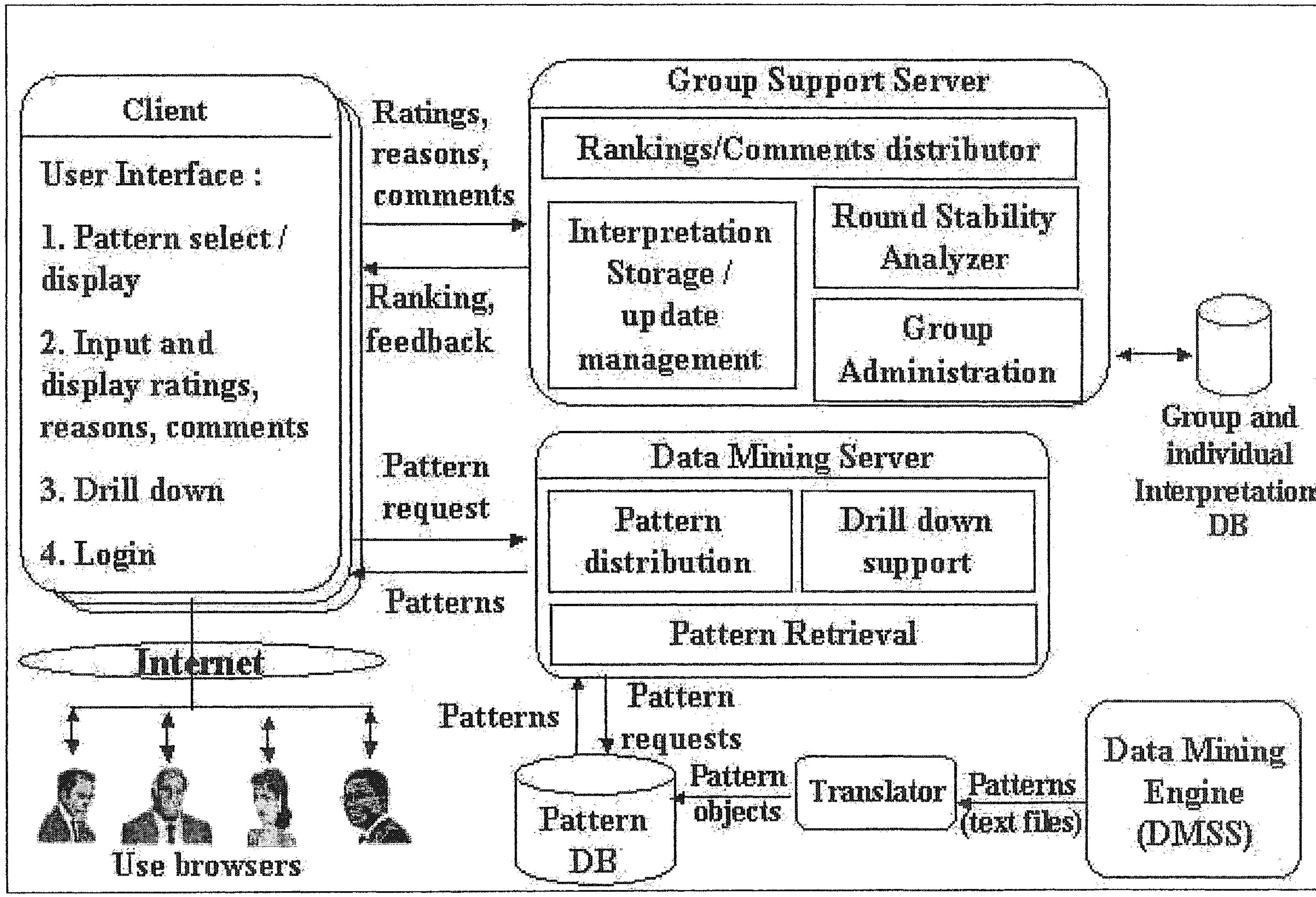


Figure 3. CIS conceptual architecture

1.5 Upcoming Chapters Preview

Chapter 2 of this dissertation reviews related research. Chapter 3 introduces the domain of infection surveillance in hospitals generally and at UABH in particular, and describes the profiles of the experts and the criterion judge of the experiment. Chapter 4 describes the methodology of this research, which includes the design and implementation of the CIS prototype, the experiment, and the establishment of the criterion judgments. Chapter 5 reports the results and the analysis. Chapter 6 provides our conclusion, and discusses opportunities for further research using the CIS and the results.

CHAPTER 2

BACKGROUND

CIS is in the class of software systems called computer-mediated communication systems. It applies the knowledge in that field to the problem of data-mining interpretation and evaluation in KDD. CIS drew on research in the fields of computer-mediated communications, expert-group decision-making, and human computer interaction. This chapter reviews related literature in those fields and the data-mining system, DMSS.

2.1 Human-centered Knowledge Discovery and Interpretation

As discussed in Chapter 1, most of the KDD research has focused on the data-mining part of the process. However, human involvement is essential. This section reviews the literature concerning human involvement and interpretation in the KDD process.

In his overview of the state of the art of KDD, Fayyad [22] acknowledged the importance of user interface and interpretation issues. Brachman [12] offers a human centered view of the complete KDD system in which a human user should be involved in each step to drive the discovery. However, the interpretation step in his model involves a single user only. Silberschatz [57] defined two types of users involved in the KDD process. One is the single end user, and the other is the KDD developer. All of these approaches assume that the use of a single evaluator is sufficient for effective interpretation.

For example, GTE's Key Findings Reporter, KEFIR, for health care data goes to the extent of generating reports in natural language and business graphics [36]. Matheus claims that, when compared to the reports created by health care consultants, the KEFIR reports were considered comparable, with additional unexpected findings deemed interesting by the users [36]. However, their study did not provide a means for the users to determine if the findings were indeed useful or any mechanism to capture their decisions and rationale.

A survey of KDD literature reveals that most KDD research has relied on the researchers themselves or a single expert to evaluate the data-mining result, once again assuming that the interpretation is so straightforward that a single expert is sufficient for effective interpretation. The following are some examples:

1. "we found it intuitive and straightforward to interpret the output … [31]." These researchers, like many others, performed the evaluation themselves and believed that their results were easy to interpret. However, such belief is questionable because their impressions are subject to bias.
2. "Because we had performed the validation process ourselves in the previous case study, we decided to use the help of a marketing expert in another seasonality analysis… [2]." It is good to see that they used a real evaluator; however, they stopped at just one. As their user pointed out to them, "different experts can arrive at different evaluation results using the same

validation process." Therefore, they should have used multiple experts for evaluation.

3. "Bridge experts are comfortable accessing and translating knowledge in the rule form [69]." Their report sounded as if they had multiple experts review their results. However, when questioned by the present author about how they collected and compiled the experts' evaluation, they revealed that they only used one expert. They pointed out that the ideal was to use multiple experts but that it would be difficult to coordinate and process the opinion.

One can add to this list by surveying the past eight KDD conference proceedings [21, 27, 28, 49, 58] and the other KDD literature [3, 23, 34, 44, 48].

Uthurusamy [66] offered a more far sighted view, which emphasized that the KDD process should support human-assisted computer discovery as well as computer-assisted human discovery. He pointed out that the trend should be towards collaborative discovery by a group of experts in a distributed environment.

Since the development of data-mining algorithms is an ongoing research process, cross-disciplinary collaboration during interpretation could help the typical users to sort out real, significant patterns from ones that are due to idiosyncrasies in reporting. Farrington et al. suggested such collaboration based on a study of *Salmonella agona* outbreak using their statistical-algorithm-based infectious disease surveillance system [19]. In their study, epidemiologists dismissed outbreaks flagged early on by the system as false positives due to inexperience in interpreting the data. Hence, they failed to initiate timely interventions. Cross-disciplinary collaboration could have prevented the dismissal.

The author believes that Uthurusamy's [66] idea of collaborative discovery in a distributed environment is missing in KDD. His idea, together with the lesson from Farrington's research, contributed to the basis of CIS, which focuses on expert collaboration to perform interpretation in a distributed environment.

2.2 Computer-mediated Communication (CMC) Technologies

Another area of research that guided the CIS research is CMC [63]. CMC is an umbrella term which covers the research and development that has been referred to by many names, such as Group Decisions Support Systems (GDSS), Groupware, Computer-supported Cooperative Work (CSCW) [9], and Distributed Group Support Systems (DGSS), to name a few. CMC functionalities can be classified into three major areas: communication, collaboration, and coordination [25]. The present research focused on computer support of collaboration for the interpretation of data-mining results.

In computer-supported collaboration, there is a high demand for a seamless intellectual integration of the parts for a specific purpose [61]. CIS was designed to provide an asynchronous mechanism to encourage uninhibited individual participation and to compile and quantify the individual contributions to represent the group's collective opinion. The mechanism provided a non-threatening environment that allowed equality of participation, facilitated decision refinement, and avoided forcing consensus.

There are many decision-making methods, and many of them have been adapted for supporting group decision-making [7, 17, 35, 50, 51]. Table 1 compares four commonly used group techniques based on six characteristics.

Table 1. Characteristics of group decision methodologies

Characteristics	Interacting Groups	Nominal Groups	Analytic Hierarchy Process (AHP)	Delphi Method
Structured	No	Yes	Yes	Yes
Asynchronous	Face-to-face meeting	Face-to-face meeting	Face-to-face meeting	Face-to-face or asynchronous
Non-threatening, equality of participation	High potential for member dominance	Member equality in idea generation	Member equality in pros-cons elicitation	Optional anonymity, equal participation
Structured decision refinement	Unstructured, may not refine ideas	Can be modified to include one iteration of refinement	Can be modified to include iterations for refinement	Structured iterations for refinement.
Not Force Consensus	Conformity pressure	Aimed at obtaining consensus	Aimed at obtaining Consensus	Record both agreements and differences
Compile and quantify opinions	Difficult	Yes	Yes	Yes

As Table 1 indicates, the technique best suited for the structured group communication process in CIS is the Delphi Method, and the reasons are discussed in the next section.

2.2.1 The Delphi Method and the CIS

The Delphi Method is an iterative process, facilitated by a moderator, to structure group communication aimed at producing detailed critical examination and discussion of issues and not at forcing a quick compromise [64]. The iterations in Delphi are called rounds. In the first round, the group members brainstorm to generate the issues to be discussed. The moderator compiles the issues to form the starting point of the second round. In the second round, the group members vote on the issues independently and submit the ratings, together with supporting reasons, to the moderator. The votes can be in the form of ratings or rankings.

In the subsequent rounds, the members are provided feedback for each issue, along with their votes from the previous round. The feedback includes the group's aggregate vote and the reasons. After examining the feedback, the members may change their votes and/or reasons for the current round. The moderator uses statistical techniques to determine if the group's votes are stable when compared to the votes of the previous round. The Delphi process ends when stability is reached. Many techniques were used in various Delphi studies to determine vote stability. The method used in this research is discussed in section 4.1.2.2.

The Delphi Method is particularly suitable for supporting the collaborative interpreting process because of the following characteristics [35, 64]:

1. Asynchronous - the group is distributed both geographically and temporally, hence an asynchronous process is essential.
2. Anonymity - the method supports anonymity but it is not a mandatory requirement. In a computer-based Delphi, various levels of anonymity can be implemented according to the group's profile and dynamics [64]. This flexibility allows the CIS to be adapted to different type of groups. For the hospital infectious disease surveillance application in this research, anonymity was used to prevent the existing power structure in the hospital from influencing a user's judgment.
3. Iterative with feedback mechanism – During the iterations, the group's collective evaluation is provided as feedback for review. The feedback allows

the group members to share and benefit from each other's expertise, and to refine the interpretations.

4. Responses are analyzed statistically - the analysis is quantitative rather than qualitative. It provides an empirical basis to determine the stability of the group's opinion.
5. Tolerates differences in opinion - Unlike many group decision making processes, the Delphi Method does not force consensus and captures both agreements and differences as the group's opinion.
6. Captures both quantitative and descriptive opinions - The group's opinions include both quantifiable data, such as ratings, and descriptive data.
7. Keeps the experts focused on the task.

Originally, Delphi was applied using a paper and pencil communication process. One of the drawbacks of the paper-based Delphi is the time required at each iteration for communication and data analysis. In the present research, computer technology has been applied at the various steps of the process. Advances in computer communication technologies offer opportunities to develop computer-based Delphi processes, which are more flexible and timely than the paper and pencil form [64]. Turoff [64] has developed EIES, a distributed group support system, which has incorporated the Delphi process. EIES is written in the Smalltalk programming language and is provided as a paid service to operate group sessions. Other Delphi software only supports a particular step of the process, e.g., gathering of group members' input, dissemination of the group's collective opinions, or computer statistical analysis of the input [35]. An embeddable or application programming interface for computer-based Delphi process, however, is not readily available. Therefore, the CIS Delphi functionalities were implemented as a domain-independent Java programming language package that could be used to build Delphi applications. It was used to develop the web-based infection surveillance application of this research.

The following are some of the differences between traditional Delphi and the CIS Delphi:

1. In traditional Delphi, the issues or trends to be evaluated are generated by the panel through brainstorming. Hence, the issues or trends are subjective, and the time required to generate them is substantial. In CIS, the brainstorming phase is replaced by data mining. The issues are the patterns generated by a data-mining system using data in databases. Hence, CIS issues are data-driven and are objective.
2. Since the CIS issues are generated by data mining, the voting starts at the first round instead of the second one.
3. The CIS automated most of the moderator's tasks such as collecting and compiling the votes, determining round stability, and communicating feedback. This allows a human moderator to focus on the content of the group members' written issues.
4. In the CIS experiment, the moderator's task was only to review the optional elaborated reasons entered by the panelists for the purpose of editing out inappropriate language. The moderator's approval is required before a reason is viewable by the panel.

Figure 4 shows the Delphi process adapted for the CIS. The CIS implementation of the Delphi Method is described in Chapter 4.

2.3 DMSS

The data-mining engine used for this research, DMSS, was originally developed at the Department of Computer and Information Sciences at the University of Alabama at Birmingham [13]. It has since been employed to detect the emergence and spread of anti-microbial resistance at a number of hospitals, including UABH. A brief description of the DMSS data mining technique is presented here. Brossette provides a more detailed description [13].

2.3.1 Association Rule and DMSS Pattern

DMSS is an exploratory data-mining process for automatically identifying new, unexpected, and potentially interesting patterns in hospital infection control and public health surveillance data. This process, and the system based upon it, uses association rules to represent outcomes, and association rule confidences to monitor changes in the incidence of those outcomes over time. Patterns of interest for DMSS consist of comparisons of temporal histories of association rules and their confidence and support measures of discrete or categorical data. An association rule is defined as "an expression X => Y, where X and Y are sets of items" [5]. X describes the left-hand side of the rule and Y describes the right-hand side. The meaning of such a rule is that database records that contain X tend to contain Y. Table 2 shows the general format of a DMSS pattern described by an association rule.

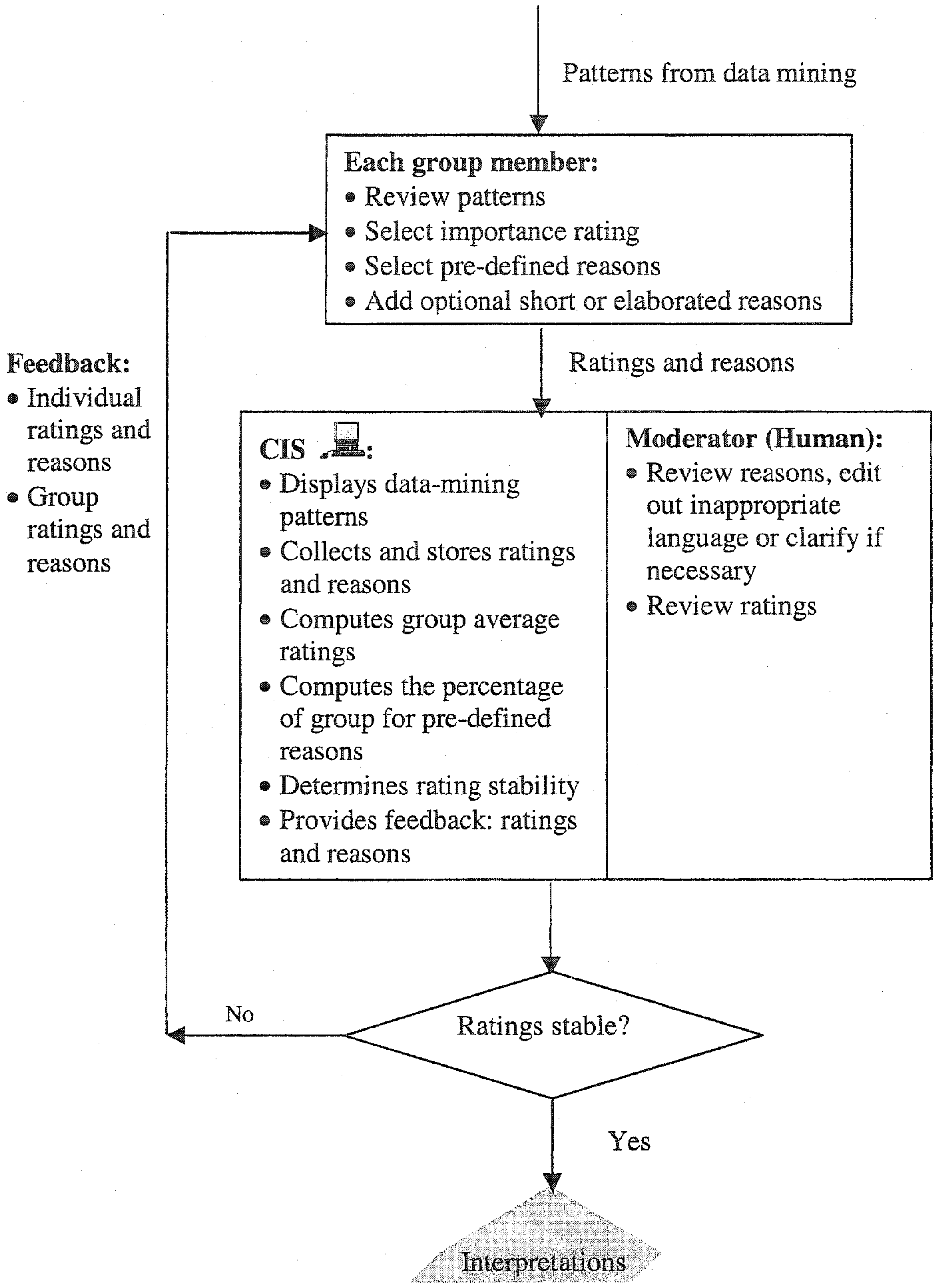


Figure 4. The CIS Delphi Process

In the infection surveillance application, the left-hand side of the rule is called the surveillance group, and the right-hand side is called the outcome. In the general format shown in Table 2, n1 and n2 are the number of records that contain both X and Y in the periods 1 and 2 respectively, and d1 and d2 are the number of records that contain X in those periods.

Table 2. General format of a DMSS association rule

Surveillance Group		Outcome	period 1	period 2
X	⇒	Y	n1/d1	n2/d2

2.3.2 The DMSS Pattern Discovery Process

Figure 5 shows a summary of the DMSS pattern discovery process.

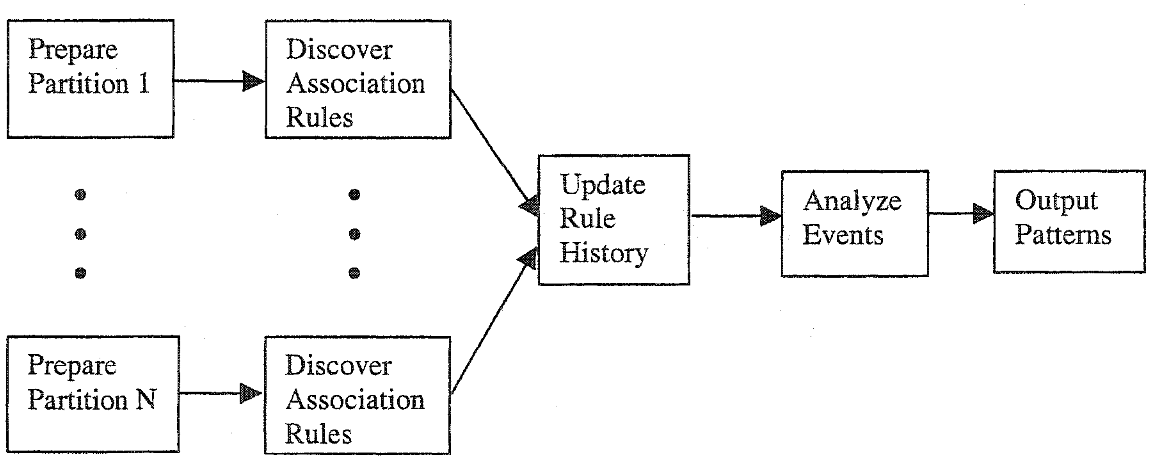


Figure 5. The DMSS pattern discovery process

The tasks of DMSS are surveillance and deviation detection. To accomplish these tasks, temporal data are divided into partitions. DMSS first discovers all association rules in each data partition. DMSS then monitors these association rules over time and detects changes or deviations in their behavior. Upon detection of a significant deviation, it triggers a pattern and reports a description of the event to the user for interpretation and verification. Table 3 shows a pattern uncovered by DMSS in August 2001, using the clinical Laboratory Information System (LIS) data at UABH.

Table 3. Example of a DMSS pattern presented as an association rule

Surveillance Group		Outcome	2/2001	3/2001	4/2001	5/2001	6/2001	7/2001	8/2001
TBIC nosocomial	⇒	A BAUMANNII R~CIPRO	3/41	0/34	0/32	1/44	0/48	1/37	5/43

This pattern means that in August 2001, 43 bacterial isolates[5] from the patients in the TBIC unit were classified as nosocomial (hospital acquired)[6]. Among these 43 isolates, 5 of them were tested positive for the bacteria A. Baumanni (Acinetobacter Baumannii/Haemolyticus), which was resistant to the antibiotic Ciprofloxacin. When the August outcome was compared to that of the previous six months, DMSS considered the increase significant and hence generated the pattern as output to alert the hospital infection control staff. The output of DMSS is a text file. Appendix A shows an example pattern from a DMSS output file and explains the data in the output pattern.

---

[5] A bacterial isolate is a test sample taken from a specimen (e.g. urine, blood) obtained from a patient.

[6] DMSS classifies a bacterial isolate as nosocomial if it is from a sample collected on or after the patient's third day in the hospital. That meant the bacteria was likely acquired in the hospital [13].

2.4 Human Computer Interface Research

As shown in Appendix A, the output from DMSS is a text file which lists patterns uncovered during a time period, such as a month. The number of patterns discovered in a month for a major medical center like UABH could range from 20 to more than 40. From Moser's preliminary retrospective study (described in section 1.1), it was clear that the lack of a simple, easy-to-use interface to access and review the output of DMSS was a significant barrier to gaining interest from the end users. Therefore, the user interface is a very important part of the CIS. Nielsen [45] found that the effort for the design and development of a usable user interface for a software system usually ranges from 20 to 60% of the total development effort. For the application in this research, the user interface development effort ranked at the high end of that estimate. Therefore, much of the research in the field of human computer interaction was used to guide the design and development of the user interfaces in CIS [8, 37, 45, 55]. To avoid cognitive overload, one of the guiding principles for the CIS user interface was "less is more" [45]. In a computer-mediated collaborative environment, the problem of providing the user with the right amount and type of tools is compounded by the differences in individuals' mental models of the technology, and the organizations' structure and culture [46]. Therefore, special efforts were made to use the proper terminology of infection surveillance and to accommodate the experts' various work styles and the constraints in the work environment. Ideas from the extensive research on group behavior using group decision support systems by Hiltz et. al. [29] were used to guide the design of the CIS tools. The user interface for this research presented the DMSS and the CIS as an integrated system called Infection Surveillance System (ISS) to the end user.

CHAPTER 3

PROBLEM DOMAIN ANALYSIS

Each year in the United States nosocomial (hospital-acquired) infections affect two million patients, costing more than $4.5 billion, and accounting for half of all major hospital complications [41]. Even more alarming, the number of drug-resistant infections has reached unprecedented levels [4]. The emergence of highly resistant bacteria has resulted in significant morbidity and mortality. Microenvironments, especially hospitals and hospital intensive care units, are foci where resistant organisms originate and propagate, spreading to larger environments as opportunity provides [30, 32, 43, 54]. Early recognition of emerging problems requires proactive surveillance as part of an infection control program at the hospital and sub-hospital levels [30, 32, 43, 54]. Such surveillance entails the examination of large volumes of hospital laboratory data. However, extensive analyses of hospital data requires considerable time and resources, both of which few hospital epidemiologists have in reserve. Consequently, these data are underutilized, and the knowledge they contain is either not discovered in a timely fashion or goes completely undiscovered.

3.1 Traditional Infection Surveillance

"Epidemiologic surveillance is the ongoing and systemic collection, analysis, and interpretation of health data in the process of describing and monitoring a health event" [40]. With few exceptions, hospital surveillance efforts are based upon predetermined criteria or events that are considered interesting and significant. The methods used for surveillance include manual review of suspected cases of nosocomial infections, ad hoc reports from staff of suspected outbreaks, the tabulation of basic summary statistics, and construction of annual or semi-annual, hospital-wide antibiogram summaries. These measures are not timely and often miss emerging, complex patterns [43]. Consequently, it has been widely recognized that active, timely, and sophisticated intra-hospital surveillance is needed [43, 54].

The main weakness of traditional surveillance methods, including existing computer-aided ones, is that none are capable of identifying patterns that are not suspected a priori. Therefore, if a problem is not suspected, it usually goes undetected. The discovery of patterns that would not be discovered by traditional analysis is a fundamental premise of data mining. The data mining technology of DMSS is especially well suited for epidemiologic surveillance. DMSS automates the extensive analysis of hospital clinical laboratory data to detect new trends of infection and drug resistant bacteria [14].

3.2 An Idealized Computerized Surveillance System Using DMSS and CIS

Dean and coworkers [16] describe a hypothetical ideal computer-based public health surveillance system that detects all unusual trends without having to provide specific directives as search constraints. The user then identifies the most interesting trends, investigates them with traditional queries into the data, and finally takes the appropriate actions based on the findings. This is essentially a description of the use of a data-mining tool for active epidemiologic surveillance. DMSS is the first generation of these tools. A weakness in Dean's model, however, is the assumption that a single expert can evaluate all novel patterns efficiently. In complex domains such as medicine, consultation and discussion among experts of the same or different disciplines is common, especially when presented with new knowledge. Such collaboration occurs within an institution, across a region, and across the globe. This is evidenced by the active discussions that can be found in medical discussion groups such as ProMED[7]. An ideal computerized surveillance system therefore should detect all unusual trends with very little direction from the user AND should allow for the timely collaboration of a diverse group of geographically and temporally dispersed experts. Therefore a combined system based on a data-mining mechanism and the CIS will satisfy these requirements.

3.3 Infectious Disease Surveillance at UABH

The UAB infectious disease surveillance process is a prototypical example of the traditional epidemiologic surveillance described in section 3.1. Data are collected at the hospital's Clinical Microbiology Laboratory, where various samples taken from patients, such as blood, urine, and tissue, are tested for presence of pathogenic bacteria. If present, these bacteria are then assessed for susceptibility or resistance to a standard set of antimicrobics. The test results are stored in the LIS that serves as the Infection Control database. This database is DMSS's data source.

By federal health regulation, the hospital has in place an Infection Control Committee (ICC). The Committee is responsible for monitoring infectious disease activities in the hospital, such as:

---

[7] ProMED's website is at http://www.fas.org/promed/

- Tracking the occurrence and frequency of infectious diseases and anti-microbial resistance behavior according to the areas in the hospital. The major focus is hospital-acquired infections, especially in the Intensive Care Units (ICU), where more than 60% of the all these infections occur.
- Tracking employee health status and safety practice policy compliance.
- Discussing and making recommendations for new policies aimed at preventing infection.
- Monitoring infection control policies compliance by hospital units.
- The committee is comprised of 26 appointed members from many disciplines throughout the hospital, including Infectious Disease, Surgery, Microbiology, Infection Control, Pharmacy, Employee Health, various areas of Nursing, Safety, Risk Management, Bio-safety, Environment Services, and Hospital Administration. The committee meets every two months. The meeting agenda usually includes updates of the latest surveillance data, employee health status, current policy and practice, and discussion of new techniques or equipment intended to prevent infection.

In order to better understand the problem domain, the author interviewed key members of the committee. The interview with the committee chair, an Infectious Disease physician, revealed that there is a subgroup of five people that meets weekly. This subgroup is comprised of the Infection Control Practitioners (ICPs), a representative from the clinical laboratory, and the chair. The purpose of the subgroup meetings is to review the surveillance data to identify problems, whereas the purpose of the whole committee meeting is for dissemination of findings to the various departments. In addition, committee members from various functions (e.g., the microbiology laboratory or pharmacy) might meet with the chair prior to the bimonthly meeting to review data and issues.

Interviews with the committee chair, head of an ICU (a physician), an infectious disease physician, a microbiologist, and an ICP elicited the following information needs:

1. To integrate the various data sources throughout the hospital to assist investigations, e.g., specific information about a hospital unit, or diagnosis and drugs used together with laboratory test result.
2. To include more information in the Infection Control database, such as patient diagnosis.
3. To gain access to timely, summarized data of bacteria sensitivity to antibiotics by hospital unit.
4. To gain access to a mechanism to facilitate the group to work together, given busy and varied schedules, travel, and dispersed locations.
5. To maintain a memory of infection monitoring profile over time so that job rotation will not cause discontinuity in the monitoring. This is because the ICPs are responsible for monitoring the infection patterns of specific units. However, they rotate to different units every three months. Hence, there is no continuity in the patterns that they monitor.
6. To gain access to the automatic discovery of microbes that show increased antibiotic resistance. The trends and the patterns of infection and antibiotic resistance are useful information that can be used to alleviate the current manual review process used to develop the antibiogram, which is the

guideline for physicians in the empirical use of antibiotics in various area of the hospital.

7. To develop a capability to benchmark with hospitals in the region.

The use of the CIS in conjunction with a data-mining tool, such as DMSS, would meet needs three through six. Need number seven is a potential application of CIS when extended for regional, national, and global epidemiologic surveillance. The global conceptual model for epidemiologic surveillance described in [68] could meet such a need.

Needs one and two involve better integration of all patient data online so that all data can be accessed via a single entry point. Although they are the functions of the information processing units in the hospital and are outside the scope of this research, this background information heightened the importance of keeping the users' expectations in line with the research as they participate in the studies of CIS.

3.4 Applying CIS and DMSS for Infectious Disease Surveillance at UABH

Figure 6 shows a model that supports collaborative interpretation of the patterns from DMSS using the CIS for infectious disease surveillance at UABH. In this model, individual users will interactively review the patterns from DMSS. The individual's goal is to use his or her domain knowledge and experience to gain understanding of the findings and to prioritize the findings with respect to epidemiologic investigation. During the first iteration, each user decides on the priority independently by assigning importance ratings to the patterns and submits the decision to the CIS. In addition, the user could submit supporting reasons to accompany the ratings. The CIS supports a human moderator to review the comments, to screen inappropriate language, and to clarify the experts' input. The CIS computes the aggregates that represent the collective interpretation of the group, which is then fed back to individuals together with their original submissions. The individual may modify his or her ratings, or add new

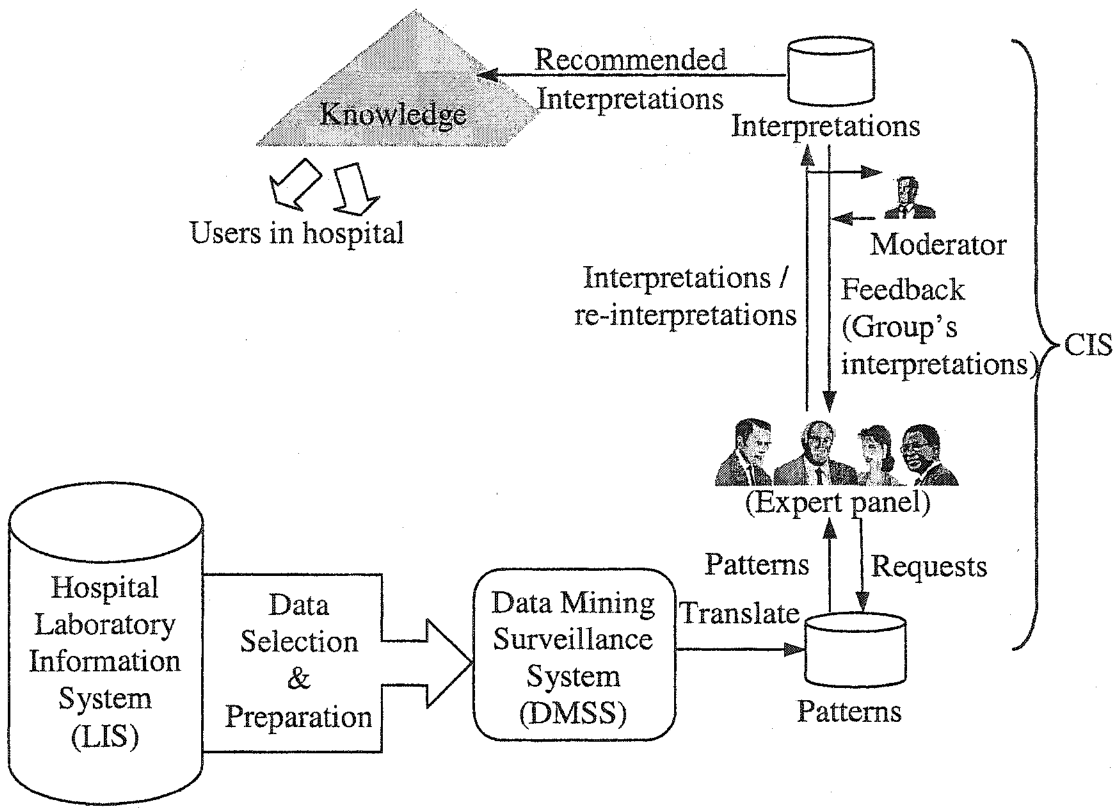


Figure 6. A Model for collaborative interpreting of DMSS output

comments after reviewing the feedback and then resubmit. The feedback and resubmit cycle may repeat for a few rounds. Participants will work asynchronously at the times and locations that are convenient to their schedules. When the CIS's analysis of the submissions indicates stability, the result will be provided to the group for the final voting of acceptance. All reasons and ratings, including the ones that deviate, are recorded so that disagreements will not be lost. The resulting interpretation is the group's recommendation of investigation priority to the infection control staff.

A typical user of the CIS is likely a member, or a designated representative by a member, of the ICC, who plays a role in making decisions that impact infectious disease control in the hospital. Table 4 summarizes the profiles of potential CIS users based on the author's observations at ICC meetings and interviews with committee representatives. The understanding of the potential user profiles was essential for the design of the user interface.

3.5 Profile Summary of the Experts and the Criterion Judge

To test the effectiveness of CIS experimentally, we used experts in the roles that indicated high interest in the CIS and data-mining technology shown in Table 4, "Profile Summary Of Potential CIS Users." The involvement of the ICPs could have been beneficial. However, the ICPs at UABH showed very little interest in the research. During their interviews with the author, they recounted a previous involvement in research. Their efforts contributed to a successful technology development but resulted in a reduction of their head count. They were still bitter about the loss of personnel and hence would not want to participate in research that they perceive could be a threat to their jobs.

In CIS, panelists could be put into one of the two categories, in-group versus out-group. All panelists would perform the same tasks except that out-group panelists would not be included in the group interpretation computation, and the in-group panelists would

Table 4. Profile summary of potential CIS users

Role	Education / Training	Work locations	Purpose of information need	Information need	Computer Skill	Interest in technology (CIS)	CIS user type
Lab Director / scientist	M.D. or Ph.D. in microbiology	UABH and affiliated hospitals	• Quality assurance • Lab performance monitoring • Communicate antibiotic resistance changes	Drugs and microbes association: • Trends • Changes	Likely skilled user of PCs, email, WWW, LIS, HIS	High to moderate	Panel expert
Clinician: Infectious Disease Specialist	M.D. with Infectious Disease Specialization	UABH and affiliated hospitals and clinics	• Monitor nosocomial infections • Track patterns of transmission • Track trends of antibiotics sensitivity • Guide strategic use of antibiotics to achieve the lowest mobility / mortality at the lowest cost	Infection information: • Type • Location of occurrences • Pattern over time	Varies. Moderate to skilled user of PCs, email, WWW, and HIS	High	• Panel expert • User of CIS output
Clinician: other practices	M.D. with specialty training	UABH and affiliated hospitals and clinics	• For effective drug use • For proactive patient follow ups • To determine epidemic vs. endemic situations	• Comprehensive, timely and integrated patient data • Antibiotics sensitivity data by hospital unit • Effectiveness of drugs	Varies. Moderate to skilled user of PCs, email, and WWW	High only if convenient	• Panel expert • User of CIS output
ICP	R.N. or former Lab Technician	UABH or nursing homes	• To monitor hospital acquired infection occurrences • To track the trends • To report to ICC and ICC chair	• Surveillance data of assigned unit • Counts of occurrences and percentages • More integrated patient data	Varies. User of hospital information systems. Might be email and WWW user.	Low. However, their participation is actually very important.	• Panel expert • User of CIS output
Pharmacist	Varies from B.S. to PharmD.	UABH wards	• For cost and effective drug use	Changes in drug sensitivity and/or resistance	User of email, WWW, Drug Utilization Information Systems.	High	• Panel expert • Operational: browse drug sensitivity.
Lab Bench Technician	B.S., M.S. in microbiology	UABH lab	Daily operational information	Aggregate and counts, patterns of drug sensitivity changes	Familiar with LIS	Low	Operational: browse drug sensitivity.

not see the opinions of the out-group panelists in the feedback. However, feedback to the out-group panelists would include the opinions of both groups.

For this experiment, there was a 6-member in-group panel and a 1-member out-group panel. The in-group was the collaborating panel in this research to produce the collaborative interpretation of the DMSS patterns. The in-group panel consisted of two infectious diseases physicians, two critical care physicians, a pharmacist, and a clinical microbiologist. One of the infectious diseases physicians was on the ICC. The other panel members were recommended by members of the ICC. The two critical care physicians were recommended by infectious diseases physicians. In general, critical care physicians have not been on the ICC; however, the chair of the ICC thought their input could be useful because of the patient types they work with. The in-group panelists were asked to provide self-assessments of their computer skills, and the result is shown in Table 5. Most of the experts' computer screen resolutions are 640 pixels by 480 pixels.

Table 5. Summary of panelists' self-assessment of computer skills

	1 (Novice)	**2**	**3**	**4**	**5 (Expert)**
General skill (keyboard, mouse)	0	0	1	1	3
Email	0	0	2	1	2
Web browsing	0	1	0	2	2
Web form filling	1	1	0	2	1

Due to subject confidentiality, the profiles of the in-group panelists are summarized in Table 6. The ages of the panelists ranged from early thirties to fifties. The panel consisted of both male and female participants.

Table 6. Profile summary of expert panel

	Infectious Diseases Physicians	Critical Care Physicians	Pharmacist	Microbiologist
Training	M.D. with Infectious diseases specialty training	M.D. with specialty training other than infectious diseases	B.S. Pharmacy with Residency in general practice and infectious diseases	Ph.D. in microbiology
Clinical care responsibilities (as % of work time)	> 50%	> 50%	50%	0%
Work schedule	Varies, weekday range from 7:00am-6:00pm. Weekends, and take calls at home.	Varies. Weekday range from 6:00am-6:pm, weekends, on call at night	Routine week day hours and one weekend per month	Flexible
Interruptions while at work	Moderate to excessive	High	Moderate to high	Low
Ease of scheduling meetings	moderate	Range from difficult to impossible	Moderate if at certain time of the work day	Not hard
Use of electronic data on the job	Medical libraries (Medline, Pubmed), PIS, LIS	Medical libraries (Medline, Pubmed), PIS, LIS	PIS, LIS, Pharmacy Bulletin, Micromedex	Libraries

The out-group panel consisted of only one expert who did not join the experiment until the second month of data. When compared to the six in-group panelists, this expert had the rare attribute of also being very knowledgeable in data mining. This provided the opportunity to observe how a data-mining-knowledgeable domain expert would perform when compared to the other domain experts. This also provided the additional opportunity to observe if the group would influence the individual.

The criterion judge was Dr. Kirk Avent, who is a professor at the University of Alabama at Birmingham, the Department of Medicine, Division of Infectious Diseases. He recently retired from a clinical appointment at UABH and as the head of Infection Control at the Birmingham Veterans Administration Hospital. Dr. Avent has thirty-four years of experience as an infectious diseases clinician and in infection control. He has a distinguished career as a physician, head of infection control of hospitals, and member of a faculty of medicine. He was honored with many awards. Dr. Avent is a skilled user of computers and information technology. He was introduced to the DMSS infection surveillance method for quite some time before he performed the criterion evaluation. He has also advised the author on the patient charts review task in this research.

CHAPTER 4

METHODOLOGY

A CIS prototype was developed to conduct the experiment needed to test the hypothesis of this research. This chapter describes the design of the CIS prototype and the experimental design using CIS. The prototype provided an interactive environment for collaborative interpreting of the patterns generated by DMSS for hospital infection surveillance. However, the group-process components of the CIS prototype were developed as a domain-independent Java package that could serve as building blocks for other group-decision-support Java applications.

4.1 CIS Prototype

The CIS is a web-based client-server system that implements the three conceptual subsystems shown in Figure 3 of Chapter 1: 1) the Group Support Server; 2) the Data-mining Server; and 3) the Interpretation Client. The following sections describe the design of these subsystems in the CIS prototype.

4.1.1 Prototype Architecture Using the J2EE platform

The CIS prototype was developed and deployed using the Java 2 Enterprise Edition (J2EE)[8] platform from Sun Microsystems. The CIS software components were

[8] Sun Microsystems's J2EE website: http://developer.java.sun.com/developer/products/j2ee

implemented using JavaServer Pages (JSP), JavaBeans, helper Java classes, and Enterprise JavaBeans (EJBs), along with Dynamic HyperText Markup Language (DHTML), and client-side Javascripts. The Java Database Connectivity (JDBC) application programming interface was used to interface to a relational database management system (RDBMS) using SQL (Structured Query Language). The RDBMS used was Cloudscape, which was bundled with the J2EE development kit.

The J2EE platform was chosen for developing the CIS prototype for the following reasons:

1. It offers an integrated platform for development and deployment of the multiple-tier web application. It integrates the various Java technologies, such as JDBC, JavaMail, and JSP into one environment. The platform allowed the developer to focus on the application rather than the low level details.
2. It manages the low level communications details, such as stubs for COBRA and IIOP/RMI. Otherwise, the network programming part would have been quite a bit of work by itself.
3. The JSPs and servlets are lightweight processes, as compared to the traditional CGI processes, which provided better performance.
4. The version from Sun Microsystem was free for educational use. In addition, it was bundled with a RDBMS, Cloudscape.

The prototype used a Model-View design. It is a simplified form of the Model-View-Controller design pattern [60]. In a Model-View design, the logic of user interface is separated from the data access and the logic that models the business rules. In the CIS prototype, such design separates the logic that modeled the Delphi process and the database access from the logic that manages the user interface.

The user interface, the view, is domain specific because it uses terminology and conventions of the medical field, especially related to infection surveillance. In addition, it is also specific to the data-mining methodology to a certain extent. For example, DMSS patterns are association rules. Therefore the CIS user interface displays the patterns as described by the rules. If a different data-mining method were used (e.g. decision trees), then the pattern display would be different.

On the other hand, the Delphi process logic, the model, is neither specific to the application domain nor to the data-mining methodology. Therefore, separating the functionalities of the user interface from the software that models the Delphi Method allows the reuse of the Delphi components for other group decision-making applications. The data storage mechanism interface, i.e., the interface to the DBMS, is also independent of the user interface and is part of the model.

The CIS Model-View design was implemented using the three-tier software architecture in the J2EE environment. Figure 7 shows the three-tier CIS architecture. The first tier is the Interpretation Client, which is a web tier application. The second tier forms the Data-mining Server and the Group Support Server. The Data-mining Server is implemented using the EJB Session Bean. It implemented the functions to access the DMSS patterns. The Group Support Server is implemented using the EJB Session Bean and Entity Beans. It implements the Delphi process and the access to the interpretation databases. The Session Bean handles the non-transactional access to the interpretation database. The Entity EJB handles the transactional access to the interpretation database.

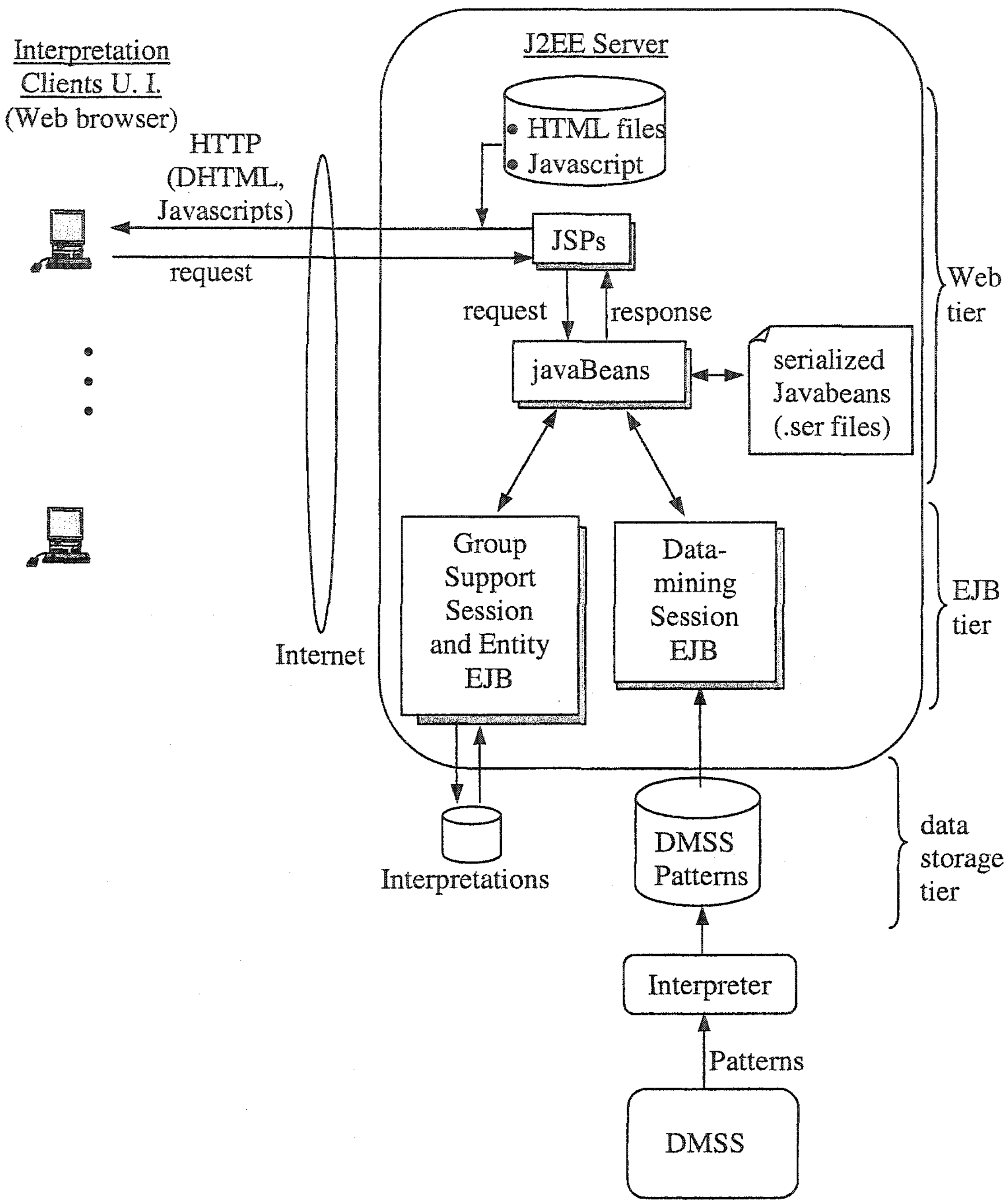


Figure 7. CIS software architecture in the J2EE environment

The third tier is the database management system for data storage. The EJBs are responsible for the communications to the database management system.

4.1.2 The Group Support Server

The Group Support Server implements the adapted Delphi method described in section 2.2.1. The core of the Group Support Server was implemented as a domain-independent Java package called groupSupport. Since the package is a collection of Java classes, it can be used in a regular Java application or a J2EE application such as the CIS. In the CIS, the Java classes in the package serve as the helper objects to the session EJB to provide the Group Support Server functions. The following are the major functions of the groupSupport package:

1. Compute the group aggregate vote as new votes are received and return the new aggregate values as result.
2. Store ongoing interpretations in the Interpretation database.
3. Determine vote stability. For the research, this function determines if the ranking of the current round has changed from the previous round. If there is no change, then vote stability is declared. The method to determine stability in CIS is described in section 4.1.2.2 below. Alternative methods for stability determination could be added to the package for other group process applications.
4. Track the round number.
5. Send email notifications to the panelists about round change.
6. Store and retrieve comments to and from the database.

7. Update comment and comment status. This function was used to support the moderator in reviewing, editing if necessary, and deciding whether to approve the comment for panel's viewing.
8. Send messages to notify the moderator when comments are received.
9. Obtain votes from Interpretation Clients.

The groupSupport package also supports two other features that are useful for group experimentation:

i. In-group / Out-group members: this feature allows a group administrator to designate whether a member is part of the group or not. An in-group member is considered part of the group, and the member's input is included in the computation of the group's aggregate interpretation. An out-group member is not part of the group. The member may or may not know that he or she is not in the group when using the CIS. The input of the out-group member is not included in the computation of the group's aggregate interpretation. However, the out-group member's input and the group's aggregates are used to create the aggregate feedback for the out-group member. Therefore, the out-group member will see the whole group's input in his or her feedback. The in-group members, however will not see the out-group member's input. Such a feature could be useful in a number of ways:
    a. To obtain input from additional experts without influencing the existing group.
    b. To observe how a user performs before deciding to add that user to the group. It can also be used in the CIS training phase of a user.

c. To observe how the group would influence the out-group member without affecting the group.

ii. Quorum level: This is the percentage of in-group members that have to participate in a round before the Delphi proceeds to the next round. This feature would allow a panel to move on if a small percentage of members are behind schedule. For the current research, the quorum was set at 100% because the panel was small.

The Group Support Server uses the groupSupport package to implement it's functions. In addition to the above groupSupport functions, the Group Support Servers implemented the following functions:

1. Accept request for group aggregate and comment data from the Interpretation Clients.
2. Return group aggregate and comment data to the Interpretation Clients.
3. Retrieve the user profile from database for presentation preferences and return such preferences to the Interpretation Clients.

The Group Support Server was implemented as a Session EJB. Therefore, it communicates with the Interpretation Client using the Internet Inter-ORB Protocol Remote Method Invoke (IIOP-RMI) of the J2EE.

4.1.2.1 CIS Collaborative Interpretation Representation

Experts provide interpretations for each of the patterns returned by DMSS. The basic content of a CIS interpretation consists of an importance rating, a set of "clickable" reasons, and free text entries for reasons that are not covered by the clickable ones. The free text entries are optional. They provide an additional channel for the experts to share questions or opinions about the patterns. The importance rating is a Likert scale of 1-5, where 1 is the least important and 5 is the most important. Figure 8 shows a pattern as it appeared in the CIS interpretation screen. The right panel is the area where the user entered the importance rating and the reasons.

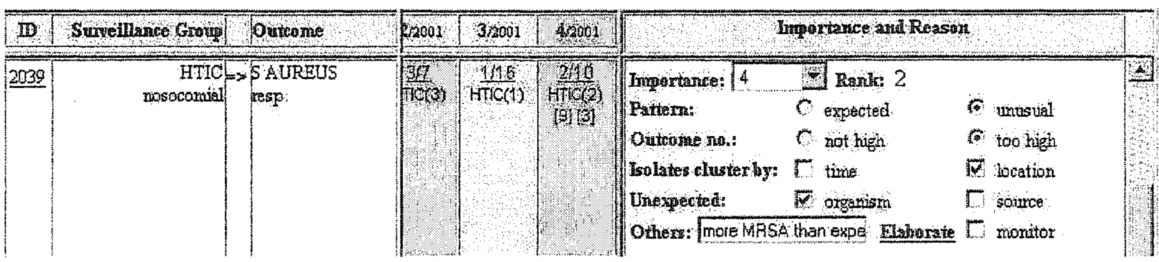


Figure 8. A pattern in the CIS user interface showing the interpretation data

There are two types of clickable reasons. The first type is a set of mutually exclusive choices that is represented as a set of related radio button choices in a Graphical User Interface (GUI). The second type is a single selectable item that is represented as check boxes in the GUI. The clickable reasons provide a set of frequently used reasons to support the judgment of a pattern's epidemiological importance. The intention is to reduce the amount of typing required of the users and hence reduce the interpretation time and promote input. The clickable reasons were designed the input from the domain experts.

There are two clickable reasons of the first type:

1. Pattern: this refers to the DMSS pattern. The choices are *expected* and *unusual*. *Expected* means that the user thinks it is not uncommon to see such kind of infection relationship in hospitals. *Unusual* means otherwise.

2. Outcome: means the number of isolates found in the outcome group of the DMSS pattern. *Not high* means the number is within the usual expectation, and *too high* means otherwise.

There are four clickable reasons of the second type:

1. Isolates cluster by *time*: means isolates reported in the pattern are from specimens around the same time frame.
2. Isolates cluster by *location*: means isolates reported in the pattern are from specimens of the same location.
3. Unexpected *organism*: means the expert usually does not expect such an organism in the kind of relationship reported by the DMSS pattern.
4. Unexpected *source*: means the expert usually does not expect the kind of relationship reported by the DMSS pattern to come from the specimen sources.

4.1.2.2 Determination of Round Stability

The strength of the Delphi Method is the ability to capture both agreements and differences in opinion. Therefore, in order to determine when to end a Delphi process, testing for round stability is more appropriate than measuring to determine if consensus is achieved. Round stability measures the stability of the panel's vote distribution curve over successive rounds of the Delphi. The stability of the vote distribution is used as the criterion for ending a Delphi process, because the interest is in the opinion of the panel rather than that of individuals.

Scheibe's histogram comparison method [53] was particularly applicable for analyzing the round stability for the small expert panel. Table 7 describes the algorithm to compute the percentage change between two successive rounds R_i and R_{i+1} using Scheibe's method of histogram comparison of vote distribution.

Table 7. Algorithm to compute the percentage change between two rounds

1. Subtract the histogram of R_{i+1} from R_i columnwise and take the absolute value of the column differences. Each column correspond to a rating unit. For the Likert scale of 1-5 used in CIS, there are 5 columns in a vote distribution.
2. Sum the absolute value of the column differences to form the total units of change.
3. Compute the net person-changes: net person-changes = total units of change / 2. This is because any one participant's change of opinion is reflected in the histogram differences by two units of change.
4. Compute the percentage change:

   percentage change = (net person-changes / number of panelists) *100

Figure 9 shows the histograms of three rounds for pattern 6299 of July 2001. The percentage changes between Rounds 1-2, and between Rounds 2-3, computed using the algorithm, are 33% and 0% respectively.

After obtaining the percentage change between two successive rounds for a pattern, the next step is to determine what represents a reasonable cut-off point at which the response may be said to be unchanged and the group opinion stable. Scheibe used observed probability of movement to compute a change level that allows the distinction to be drawn between expected "oscillatory movement" versus strong group movements that represent real changing opinion [53]. Oscillatory movements are noise. The change level represents the expected base oscillatory movement. Scheibe determined a 15% change level to represent a state of equilibrium when a rating scale of 1-10 was used [53]. This change level was used as a threshold to compare with the percentage change between two rounds for a pattern. A percentage change that is smaller than the threshold indicates that the aggregate vote distribution is considered stable, and the Delphi iteration ends. CIS used a rating scale of 1-5, therefore a more conservative change level of 7% was used as the threshold. For pattern 6299 of Figure 9, the group's opinion on the importance of the pattern is considered stable after Round 3 because the change percentage between Rounds 2-3 is 0%.

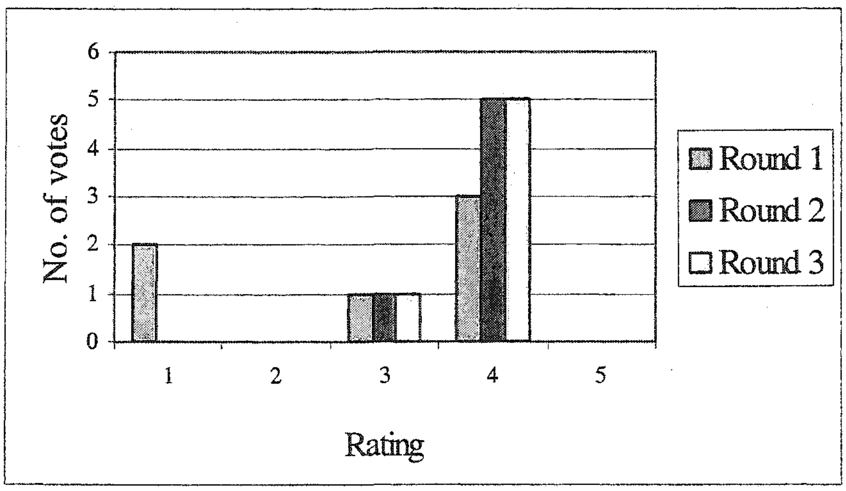


Figure 9. Pattern 6299 July 2001 ratings distribution

For this research, the stability of the importance rating for the set of patterns for a month is of interest rather than the individual patterns. This is because it is not necessary to subject the panel to another round just because a few patterns cannot reach stability. Therefore, in CIS's last step of round stability computation, the change percentage of all of the patterns of the month are aggregated and then divided by the total number of patterns to obtain an average percentage change for the pattern set between two successive rounds. The average percentage change is then compared to the threshold of 7.5% to determine round stability. Table 8 shows the modified algorithm used in CIS.

Table 8. Modified algorithm to determine the round stability of a pattern set

1. For each pattern in the set:
    i. Subtract the histogram of R_{i+1} from R_i columnwise and take the absolute value of the column differences.
    ii. Compute: total units of change = sum of absolute value of all column differences
    iii. Compute: net person-changes = total units of change / 2
    iv. Compute: percentage change = (net person-changes / number of panelists) *100
2. Compute: average percentage change = Sum of percentage change of all of the patterns in the set / number of patterns
3. Compare the average percentage change to the threshold to determine stability

4.1.3 The Data-mining Server

The major functions of the Data-mining Server are as follows:

- Accept pattern requests from the web tier Interpretation Client.
- Retrieve the requested data from the patterns database.
- Return the patterns to the client.

The Data-mining Server is being implemented as a session EJB responsible for composing SQL queries to search the patterns database based on the requests from the clients.

4.1.4 The Interpretation Client

The Interpretation Client subsystem implemented the display functions of the user interface and the logic to support user actions.

The Interpretation Client functions include the following:

1. Communicate with the Data-mining Server to obtain the DMSS patterns.
2. Generate the DHTML to display the current month's patterns and the Delphi feedback data.
3. Collect the user's interpretation in the form of ratings and reasons. The ratings and reasons are saved as work in progress on the J2EE server until the user clicks at the Submit button to submit the interpretation.
4. Communicate with the Group Support Server to submit a user's input and to obtain the group's decision
5. Save a user's state of usage so that if the user left in the middle of a session, the user state is restored when the user logs in again at a later time. This is

necessary because most of the panelists have a high level of interruption on their jobs. This function requires that every user input (a mouse click or keystroke) is sent to the Interpretation Client and saved.

6. Supports a monitor function. This function allows a user to indicate that a pattern should be monitored in the next month. When the next month's patterns are displayed the first time to the user, the Interpretation client checks if the monitored pattern is in the new pattern set. If it is, then the monitor check box of the pattern will be set to alert the user of the recurrence of the pattern.
7. Sends a notification or alert using JavaMail to the developer, the author, regarding any exceptions detected. This function allows the developer to take action about the exceptions as soon as they occur.
8. Uses error JSPs to display a friendly message to the user in the event an error happened on the web tier of the application. The error JSP will also invoke the notification function to notify the developer of the actual exception.
9. Logs the usage of some user interface features by each user. The features logged were drill pattern history, drill patient isolate listing and antibiograms, and viewing of elaborated reasons. The logs are useful for determining which features are useful.
10. Keeps records of logins and submissions. These records are used to determine the time to decision for the efficiency evaluation of the CIS. The records are stored in the database.

11. Stores the users' estimates on time spent in performing the interpretation. Each time a user submits ratings and reasons using the submit button, the user interface asks the user for an estimate of time spent in the session before submit. A comparison of such estimates to the data collected in Function 10 above may be useful in determining the user's perception of the demand of the system on their time.
12. Maintains users' profiles in a database. The profiles include the role of the user in the domain, such as pharmacist or infectious diseases physician.

The Interpretation Client is of a set of software components that runs in a J2EE server. These components are implemented as JSPs, JavaBeans, and Java helper classes. The JSP and JavaBeans together serve the user's requests and dynamically generate the response documents in the form of DHTML. JSP was used to encode the user interface layout using DHTML and Java scriptlets. JSP was used because it allowed the developer to see the flow of the final DHTML page. In the author's opinion, they were easier to write and modify than servlets for presentation heavy applications. JSPs were compiled into Java servlets by the J2EE container at execution time.

The JavaBeans and the helper classes communicate with the Group Support Server for requests of retrieval and storage of interpretation data, and with the Data-mining Server for the retrieval of the DMSS patterns. They serve as the barrier to separate the model logic from the user interface logic. The Interpretation Client user interface prototype is described in the next section. The Interpretation Client generates the DHTML. The DHTML documents are sent to the web browser using the HTTP protocol. The DHTML documents use client-side JavaScript to perform local checks and allow users to control the display of these data. The user functions supported by the Interpretation Client are described in the next section.

4.1.5 User Interface Design

The DHTML generated by the Interpretation Client presents the CIS user interface in the web browsers Netscape 4.0+ or Microsoft Internet Explorer 5.0+. Java applets and Java client applications were not considered as implementation options for the development of the user interface because the users had the option of using any of the computers on the hospital floors to participate in the experiment. The research team was not allowed to install applications on these computers, including plug-ins that would enhance the performance of Java applets. Therefore, the user interface must be viewable by using the web browsers only.

To create a coherent application front end for the users, the web user interface is labeled as the "Infection Surveillance System" on the screen. As far as the users are concerned, it is a system that presents discovered infection patterns and conducts the Delphi process to obtain their opinion, individually and as a group, on the priority of these patterns. The following principles guided the CIS user interface design:

1. Require minimum number of clicks to get to the data.
2. Minimize typed input. The purpose is to prevent user errors and to reduce user's time required to use the system.
3. Use similar screen layout for every round of Delphi, with the group feedback data displayed on a screen similar to that of the first round. This would

eliminate the individual and group-work transition problem reported by Grudin [26].

4. Adopt the reflexive perspective of collaboration in the design of the CIS. The reflexive perspective emphasizes the importance of catering to the individual's preferences in the collaborative environment [15]. The CIS achieved this by providing a simple search function within a pattern set, pattern sort functions, user-selectable profile parameters to indicate display preferences, and print functions to support users who prefer to work offline using hard copy.

The best way to explain the functions of the CIS functions is to review the user interface. The CIS user interface for the expert panel consists of four screens. In addition, a web-based user interface for the researcher to monitor the progress of the Delphi process was also developed. An example screen of the researcher's user interface is shown in Appendix C. The four screens that forms the expert panel's user interface are explained below.

4.1.5.1 Interpretation Screen

Figure 10 shows the screen for interpretation. Explanation for each of the fields follows.

1. Attributes list boxes: lists the attributes found in this month's pattern descriptions. Each list box represents a category of attributes (e.g. Microbes). Clicking on the items in the list displays the patterns with the selected attributes in either the surveillance group or in the outcome description. Multiple attributes selected in a list box have the

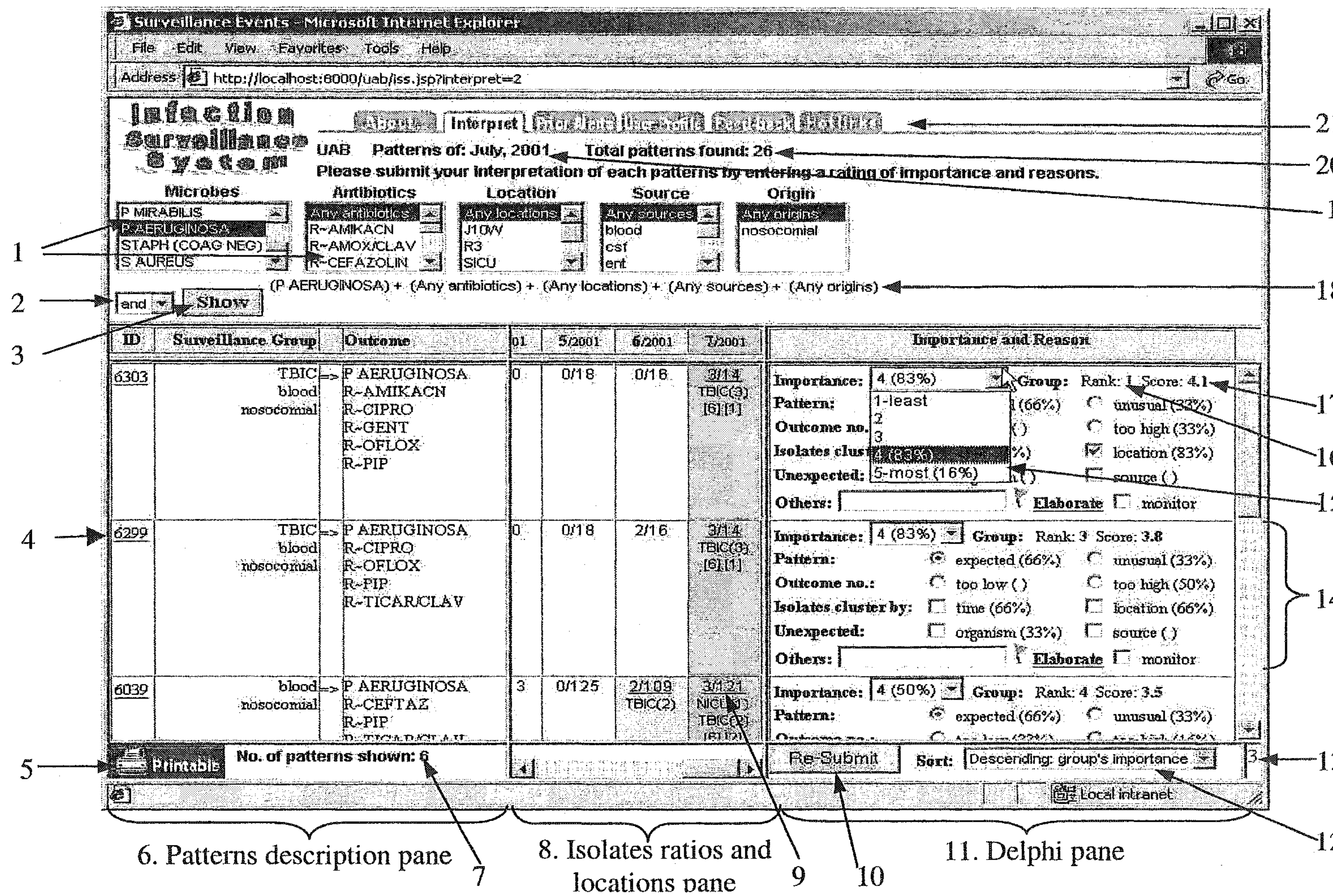


Figure 10. CIS main screen for interpretation

logical OR (/\) relationship. Attributes selected in different list boxes will have the logical AND (+) or OR(/\) relationship based on the setting of the selection in 2.

2. The between attribute category logical operator. User can select AND(+) or OR(/\) to constrain the patterns to be displayed in the bottom half of the browser window.
3. Show button: click to show only the patterns with the selected attributes according to the chosen between attribute category logical operators.
4. Pattern identification number (pattern id): a CIS-assigned number to uniquely identify a pattern described by a specific set of surveillance group and outcome attributes. The same pattern ID is used when the pattern appears in other months. The pattern ID field is a clickable link to display the history of the pattern in the pattern database.
5. Printable button: click to display a printer-friendly version of the screen.
6. Patterns description pane: this pane displays the patterns description. It consists of four columns, described from left to right as follows: Column 1 lists a unique ID for each pattern. Clicking at a pattern ID displays the history of the pattern. Column 2 lists the attributes that describe the characteristics of the surveillance group of a pattern. Column 3 is the association rule symbol (=>) that indicates the relationship of the outcome to the surveillance group. It means that in the set of isolates described by the surveillance group attributes, they also have the attributes described in the outcome pattern. Column 4 describes the attributes of the outcome pattern.
7. Actual number of patterns displayed: shows the number of patterns that meet the criteria selected in the attribute boxes described in 1. For example, in Figure 7, the user selected to display patterns with the microbe P. AERUGINOSA only. Therefore, the number of patterns displayed was 6.

8. Isolate ratios and locations pane: this pane displays the outcome/surveillance group isolates per month in the form of a ratio. Putting the mouse over the ratio displays an explanation of the ratio in the status line at the bottom left of the browser window. Clicking at an isolate ratio displays the detailed patient antibiogram for each isolates in the outcome group of the corresponding pattern. Under each clickable ratio is a list of the locations where the outcome isolates are collected, e.g., MICU(3) means three of the outcome isolates of the month is collected from MICU patients. For the ratios of the current month, the data analysis window is also displayed. They are the numbers in [ ] [ ] following the list of locations (e.g. [9] [3] means the trend of the most recent three months' isolates were compared to the nine-month trend prior to the three recent months). The color-coding scheme used in this pane is as follows: Deep pink is the month that a pattern is discovered. Light pinks are the most recent months that support the discovery of the pattern. The light yellow months are the prior months that do not cause or support a pattern to be generated. Deep blue are patterns of improvements. Light blues are the supporting months.
9. A outcome / surveillance-group ratio: This ratio indicates the number of outcome isolates among the number of surveillance-group isolates. The ratio is a link to the listing that shows detailed results of the tests on the isolates.
10. Submit / Re-submit button: use this button to submit the user's importance ratings and reasons to be compiled with the group's. When the user clicks at this button, the system also asks the user to enter an estimate of the amount of time the user has spent reviewing the patterns and entering judgments during that session. In the first round, this button is labeled as Submit. It is labeled as Re-submit in the subsequent rounds.

When this button is clicked, the system asks the user to submit an estimate of the time spent on the session.

11. The Delphi pane: this is the pane where the user enters his or her importance ratings and reasons. There are four pre-defined categories of reason that the user may click to select. The choices for the categories of "Pattern" and "Outcome no." are mutually exclusive, and hence are represented by radio buttons. The choices for the "Cluster by" and "Unexpected" categories are not mutually exclusive, and hence the users were asked to check as many as he or she deem applicable. To minimize the user's work, the Importance score is set at a default value of 1 (least important), and the default reason is Expected Pattern.

    The user may also enter short reasons in the "Others" textbox and press the "Enter" key to save each short reason entered. The user may elaborate on his or her reasons by clicking the "Elaborate" link for a larger textbox. Starting Round 2 of the Delphi process, the user may also view other panelists' written reasons by clicking this link. A flag is displayed to the left of "Elaborate" if there are user-entered short or elaborated reasons. The flag is red if the user has not viewed the new reasons associated with this pattern. The flag is green if the user has already reviewed the latest reasons.

    The "monitor" check box is an experimental feature that allows the user to indicate that the user would like the system to notify him or her if the same pattern is discovered in the future.

12. Sort selection list: The user selects a choice from this list to display the patterns in the order that the user prefers. In Round 1, the default order is by pattern ID. From

Round 2 onward, the default is by the group's ranking in ascending order. A group rank of 1 means the pattern is the most important. The possible sorting orders are:

a. Descending on group's importance ranking;
b. Ascending on group's importance ranking;
c. Descending on the user's importance ranking;
d. Ascending on the user's importance ranking;
e. Ascending on pattern identification numbers.

13. Round indicator: indicates the current round number in the Delphi process.
14. A row across the three pane corresponds to a pattern's information that is described in items 5 – 7.
15. Group's rating distribution: The importance rating drop-down list displays the percentage of panelists who voted for each rating from 1 to 5.
16. The pattern's ranking based on the group's current average importance rating. The most important pattern is assigned the rank of 1.
17. The group's current average importance rating for the pattern.
18. Summarizes the list of selected attributes and the logical relationship between them.
19. The month that the pattern set was generated.
20. The total number of infection patterns generated by DMSS in this month.
21. Navigation bar: the most important tabs for this experiment are "Interpret" and "Feedback." The "Prior Alert" tab allows the user to view prior months' patterns in the pattern database. The user interface of the "Prior Alert" tab is similar to that of the "Interpret" tab except that in the "Prior Alert" user interface, the user may choose the range of months for which to display the DMSS patterns.

4.1.5.2 *Elaborated Reason Screen*

Figure 11 shows the user interface for the input and review of elaborated reasons.

Explanation of the fields of the elaborated reason screen:

1. Description of the pattern.
2. Reasons submitted by the group members are listed here.
3. Text box for entering elaborated reasons,
4. Submit button: click to submit the elaborated reason entered in the text box.

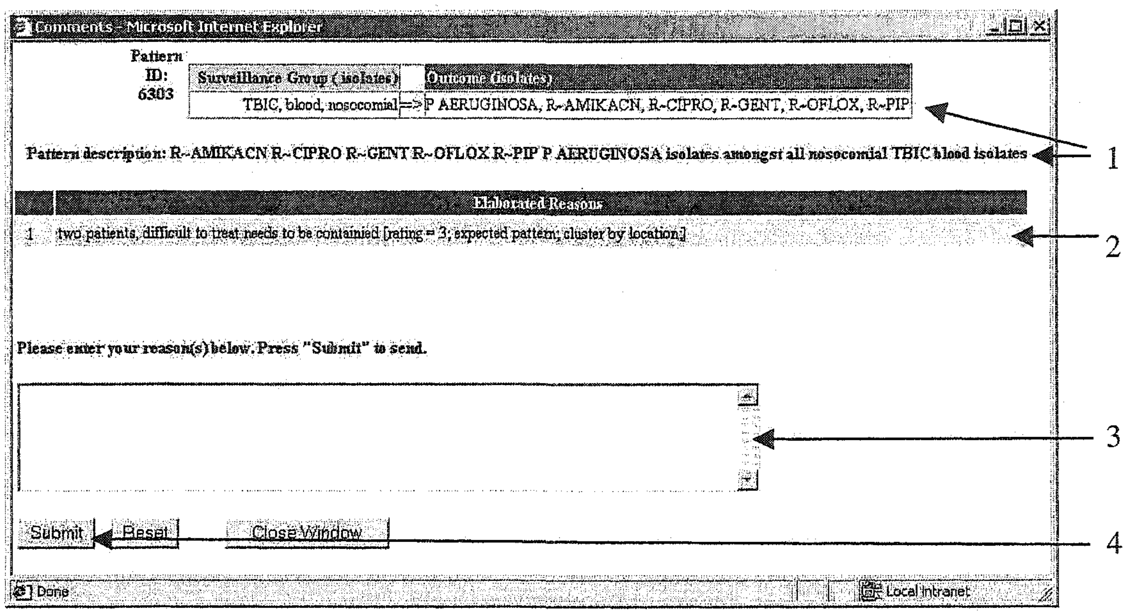


Figure 11. Elaborated reason user interface

4.1.5.3 *Patient Isolate Test Result Screen*

The purpose of this screen is to display the test results of the patient specimens that caused the generation of the DMSS patterns. Figure 12 shows the Patient Isolate Test Result Screen. This screen is displayed when the user clicks at an outcome/surveillance-group ratio in the Interpretation screen. The patient isolate test result listing corresponds to the time period of the ratio, and the pattern is displayed in this screen. Explanation of the fields in the Patient Isolate Test Result Screen:

1. Pattern description.
2. Month of the specimens.
3. Information of a primary isolate from a patient specimen.
4. Information of duplicate isolate from a patient specimen. DMSS used the following criteria for determining duplicate isolates: 1 isolate per patient per organism within 30 days with susceptibility pattern or phenotype plus or minus 2 drug changes.

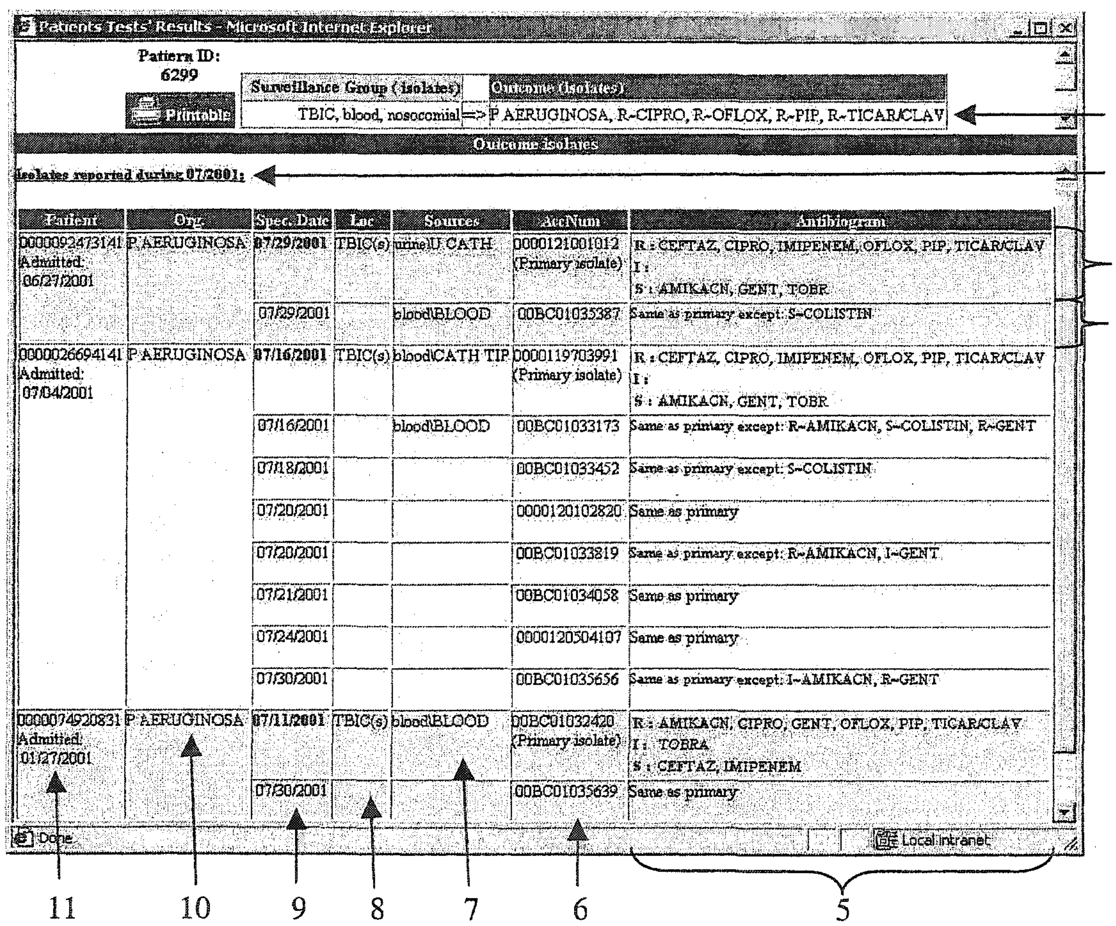


Figure 12. Patient isolate test result screen

5. Antibiograms: this is the test result. An antibiogram indicates whether the organism is susceptible (S~), resistant (R~) or intermediate (I~) to an antibiotic. For duplicate isolates, the antibiograms only contain the results that are different from the primary isolate.
6. AccNumber: specimen identification number.
7. Source: source of the specimen from the patient.
8. Location: the hospital location of the patient when the specimen was taken.
9. Specimen date: the date that the specimen was taken.
10. Organism: the organism found in the isolate.
11. Patient code and admit date: patient code is an internal identification assigned to each patient for the purpose of this study. It is not the hospital's patient identification number. The admit date is the date that the patient was admitted to the location where the specimen was taken.

4.1.5.4 Pattern History Screen

The purpose of the Pattern History Screen is to show all of the events related to a pattern over time. This display provides a graphical way to visualize the history of the pattern. This screen is useful if there are many years of data in the pattern database. Figure 13 shows the Pattern History Screen.

Explanation of the field in the Pattern History Screen:

1. Pattern description.
2. Patten history timeline. The periods in the timeline are indicated by month and year.

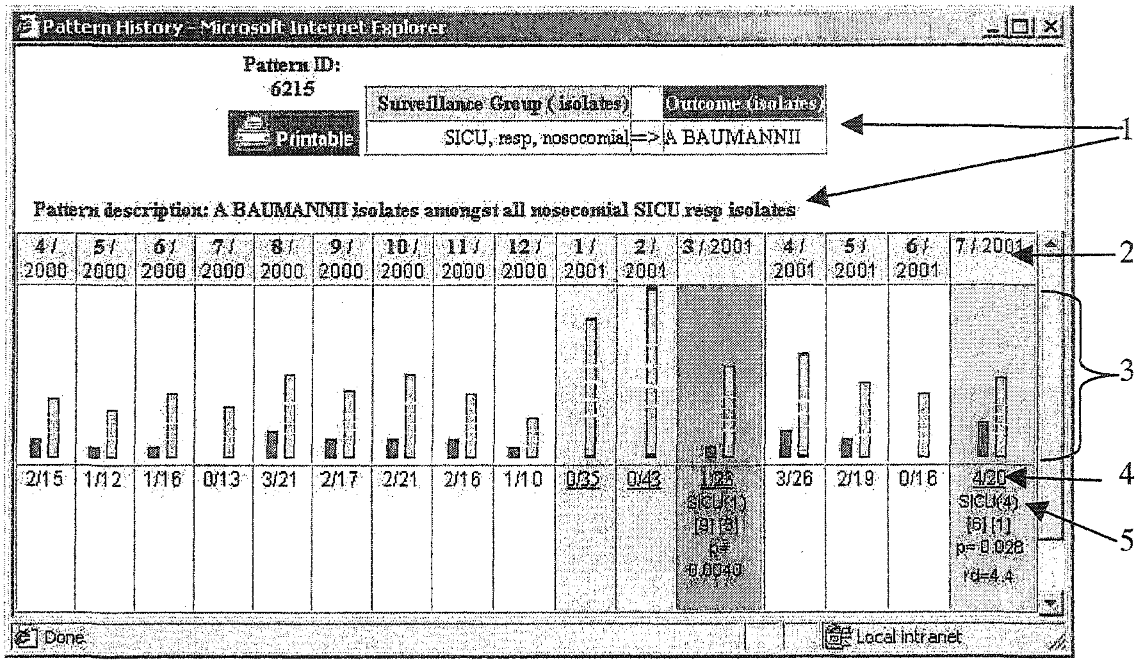


Figure 13. Pattern History Screen

3. Histogram of outcome and surveillance isolate counts. Within a time period, e.g. July 2001, the left-hand bar is the outcome isolate count and is color coded red. The right-hand bar is the surveillance isolate count and is color coded cyan. A broken bar means that the bar was scaled to fit into the display.
4. Outcome / Surveillance count ratio: this ratio represents the same information as the histogram. It is a clickable link to the corresponding Patient Test Result screen.
5. Location summary – summarizes the locations of the patients where the outcome isolate specimens were taken.

4.1.6 Interface to DMSS

The output of DMSS is a text file that contains listings of patterns. A sample pattern listing is in Appendix 1. An interpreter was written in Java to parse the text file and store the pattern and isolate data into the pattern database. In a DMSS text file, a thread ID identifies a pattern within a month's data. However, when the same pattern appears again in future months, the thread ID may be different from the one assigned for this month. Therefore, thread ID is not a unique identifier for a pattern that occurs in more than one month. However, in the pattern database, each pattern is assigned a unique CIS pattern identification number that serves as the primary key of a pattern record in the database. A DMSS pattern is uniquely described by the attributes of its surveillance group and outcome isolates. If a pattern with the same attribute description appears in a future month, the new data will be associated with the same pattern ID so that a history of the pattern is maintained. The interpreter performs the matching of the pattern descriptions.

4.1.7 The Database

A relational database is used to store the input and output of the CIS. The database for the research is local to the J2EE server environment. A number of database tables are used to manage the data in this research. The database tables were designed to accommodate the uncertainty of the DMSS data. A brief description of each type of database table follows.

4.1.7.1 Pattern Tables

These database tables contain DMSS output. This data set includes the pattern definitions, descriptions of the attributes being monitored, the pattern history, encoded patient reference number for each pattern, and DMSS analysis parameters. If additional data elements are required for drill down capability, they will be added after further analysis.

4.1.7.2 Group Vote Tables

These database tables contain the group's ongoing interpretation of the patterns. They store the ongoing intermediate output of the collaborative interpreting process. The content of the database forms an audit trial of the ratings and reasons submitted by the panel. It includes the group's aggregate interpretations generated by the Group Support Server. The composition of a group interpretation object for a given pattern is as follows:

1. The pattern identification number.
2. The date the pattern was generated.
3. The aggregate vote of the group for importance rating, and the clickable reasons.
4. The round number.
5. Vote timestamp.
6. Stability indicator (True means group vote is stable).

4.1.7.3 Individual Vote and Data Usage Tables

These database tables contain an individual's ongoing interpretation of the patterns and data about his or her usage of the functions presented by the CIS user interface. The content of the database forms an audit trial of the ratings and reasons submitted by an individual. The composition of an individual's interpretation object for a given pattern is as follows:

1. The pattern identification number.
2. The expert's identification number.
3. The expert's votes.
4. Round number.
5. Vote submit timestamp.
6. The number of times the pattern's patient reference links are used.
7. The number of times the pattern history is viewed.

The individual vote tables and the group vote tables provide the data for the study of the effect of the Delphi process on the individuals' and group's interpretations. During the experiment, these data were analyzed on a continuous basis to identify problems, and to determine the group's responsiveness and the usefulness of the drill down of patient references and pattern history.

4.1.8 Patient Data Security in CIS

CIS used J2EE's web tier default authentication mechanism. Users are required to login with a valid user name and password. The user login names and passwords were set up using J2EE's realmTool. The J2EE authentication mechanism verifies the login data and grants access for valid users to retrieve patterns and participate in the Delphi process. In addition, patient medical numbers were not used, in order to prevent identification of individual patients. Instead a patient code for the research was used to identify the patient test results from the hospital LIS.

4.2 The Experiment

The purpose of the experiment is to examine the hypothesis that the results of collaboration are better than the results that can be obtained by any single member of the group. The experiment was conducted using the CIS prototype and the patterns generated by DMSS. Retrospective data from the UABH LIS from June 2000 to October 2001 was used to generate DMSS patterns. UAB Internal Review Board (IRB) approval for using human subjects was obtained before the experiment began. An image of the IRB approval is in Appendix E. The experts were the human subjects of the experiment.

4.2.1 The Subjects

The subjects were the panel of six multi-disciplinary experts from UAB Hospital: two infectious disease physicians, two critical care physicians, one pharmacist, and one clinical microbiologist. Section 3.5 describes the profile summary of the experts. The panel was anonymous to prevent the existing power structure at UABH to influence decisions. All of the members on the panel were set up as in-group members. They were the collaborating members and were the focus of this experiment.

In addition to the six-member panel, one expert was added as an out-group member in the second month of the experiment. As described in section 4.1.2, Group Support Server, the out-group member's input is not included in the computation of the group aggregate for the experiment. Therefore, the in-group panel was not influenced by the out-group member. However, the feedback to the out-group member would include his or her input together with those of the in-group panel.

4.2.2 The Panel's Task

The task of an expert was to use the CIS user interface to collaborate with the panel to judge the importance of the DMSS patterns with regard to hospital epidemiology investigation priorities. The goal for the panel was to identify the most important patterns to investigate, assuming that there were limited resources for investigation. The experts do so by assigning importance rating to the DMSS patterns and optionally provide reasons to justify their judgments. The CIS computed the group's average rating for each pattern to determine which patterns the group considered important. The panel was instructed to use the following meaning for the importance ratings:

1 = least important (basically, the expert does not care to see such pattern in the future).

2 = not important for now (do not waste resources to investigate).

3 = mildly important (if there are unused resources, maybe follow up on it).

4 = important (not necessarily immediately, but should consider as an action item soon).

5 = very important (immediate action).

4.2.3 Experiment Design

The CIS and the DMSS patterns were used for conducting the experiment. The experts were provided individual training by the author before the experiment began. The training sessions ranged from 30 to 40 minutes. The subjects were informed that the CIS Delphi process for a set of patterns is expected to continue for a maximum of four rounds. The estimated time needed for each round is about 45 minutes. Therefore, the potential total time of participation each month is 3 hours (4 x 45 minutes).

The duration of the study was 5 months. The patterns from April to August 2001 were used as the 5 pattern sets. The number of patterns per month ranged from 22 to 40. The total number of pattern for the 5 months was 146. In the 5-month study, the panel members were presented a set of patterns each month for evaluation. The following procedure was used each month:

1. Send email to notify the panel that the patterns are ready for review and the Delphi process begins.
2. In the first round, the experts rate the patterns independently and submit their importance ratings and optional reasons.
3. Once the whole panel has submitted their ratings, the CIS computes the group average rating and determines the ranking of the patterns based on the average importance ratings.
4. CIS sends out email to notify the panel that the next round starts.
5. Experts re-examine the patterns and review the feedback to decide if they want to change their ratings. The feedback includes the group's average rating, distribution of the ratings, the ranking based on the average rating, and reasons.
6. Experts re-submit their ratings and/or new reasons.

7. Once the whole panel has submitted their ratings, the CIS computes the group's aggregate ratings and performs the stability test that is described in 4.1.2.2. If stability has not been reached and the round number is less than the maximum of 4, then steps 4-7 are repeated.
8. If stability has been reached or four rounds have gone by, start the approval round. In the approval round, the group aggregates and the individual ratings and reasons are displayed, but the user is no longer allowed to change those values. The panel members are to review the patterns' ranking, with focus on the five most important ones, and to submit their decision to approve or disapprove the final ranking. They do so by clicking at a thumb-up or thumb-down icon.
9. Once the approval round is completed, the Delphi process ends.

The following CIS parameters were used for this experiment:

i. Quorum level was set at 100%. That means all of the panel members must submit their ratings before the group could proceed to the next round.

ii. The maximum number of Delphi rounds is four.

The experts agreed to review the patterns and submit their interpretation using the CIS within a couple of days. If a user does not participate in a timely manner, the author contacts the expert by e-mail, phone call or office visit to determine if the user has decided to withdraw, or if he or she has experienced a problem with the system.

4.2.4 *Evaluation Design*

To test the hypothesis, the individual's first round ratings, and the group's final round ratings, were compared pairwise to a criterion rating.

Besides the usage log data captured by CIS, a mid-study survey and a post-study survey were conducted to evaluate the usability of the CIS user interface and the usage of the feedback data.

4.2.5 *Establishment of Criterion Rating*

The criterion rating was established by the criterion judge Dr. Avent. In addition to the DMSS patterns, Dr. Avent was supplied with relevant patient chart data for all of the patients whose specimens test results were the source of the DMSS data. A total of 240 patient charts were reviewed. The following patient chart data were collected through extensive patient charts review by experienced reviewers:

1. Reason for culture.
2. Fever status, including temperature.
3. Admit diagnosis.
4. Final diagnosis.
5. Final status (discharge or expire) and the corresponding date.
6. Method of culture taken.
7. Result of testing of culture.
8. Location history of the patient since admit date or up to thirty days prior to specimen date, whichever is more recent. Location history consists of the

name of the location, start date and end date that the patient was at the location, and the bed number.

9. Procedure history of the patient since admit date or up to thirty days prior to specimen date, whichever is more recent. Procedure history consists of date and duration of procedure, name of the procedure, and the ordering physician or surgeon.
10. Antibiotics history of the patient since admit date or up to thirty days prior to specimen date, whichever is more recent. Antibiotics history consists of the name of antibiotic, start date and end date that the patient was given the drug, method of delivery (by mouth, intravenous or inhalation), and the ordering physician.
11. Whether the patient is a Cystic Fibrosis patient.
12. Patient's symptoms related to a potential nosocomial infection according to the Centers for Disease Control (CDC)'s National Nosocomial Infections Surveillance (NNIS) System criteria [18].
13. Gender.
14. Race.
15. Age.

Dr. Avent used a user interface similar to the one used by the experts to review the patterns and enter his ratings and reasons. However, the user interface that he used contained the links to the patient chart summaries for each DMSS pattern. Therefore, he was able to investigate whether an outbreak was likely.

The establishment of a criterion rating using extensive patient chart data was possible because retrospective data was used for the experiment and research funding was available. If DMSS and CIS were used for operational infection control, such extensive patient chart review would be both cost and resource prohibitive.

CHAPTER 5

RESULT AND ANALYSES

In this research, the ratings of the panel of judges, both individually and as a group, were compared to the ratings of a single, independent judge who served as the criterion measure for evaluating the importance of the patterns. This criterion judge, Dr. Avent, was considered the "gold standard" for interpretation and evaluation for three reasons. First was his training and experience. He has had a long and distinguished career in the field of infectious diseases with over 30 years experience in hospital infection control. Both his experience in the medical field in general and in infectious disease control in particular were greater than any of the judges on the panel. The second reason his interpretation was considered the criterion was that, unlike the panel of judges, he could review relevant patient chart data in making his ratings, which gave him a decided advantage in interpreting the patterns. Third, he was familiar with the DMSS concept for quite some time before he performed the evaluation. Therefore, by his experience, access to information, and knowledge about DMSS, this judge was considered the "gold standard."

5.1 The Data

The interpretation data collected in the experiment are shown in Appendix B and Appendix C. Appendix B shows the five months of ratings data. For all of the five months, the group's ratings reached stability after round three, using the stability determination algorithm described in section 4.1.2.2. For the discussion of the results, the experts are assigned the identification of E1 to E7. The criterion judge is denoted by "C" and the group is denoted by "G". A round number notation of rn where n is the round number is appended to the identification notations to indicate which round of rating is under discussion. Appendix C shows an example of the researcher web user interface that displays the ratings of all rounds in histograms, all of the elaborated reasons entered by the experts, and summaries of the clickable reasons as percentages of experts who selected them.

To test the hypothesis that group collaboration would lead to more effective interpretations of patterns than those possible by any individual, the data were analyzed with four techniques: 1) Kendall's Correlation Coefficient of Rankings [56]; 2) an analysis using the sum of the differences of ratings between panelists and the criterion judge; 3) a graphical representation of the similarity of ratings between the panelists and the criterion judge; and 4) an analysis of hit rate and precision. Each method augments the others in gaining an overall understanding of the level of agreement between the panelists and the criterion judge.

The ratings of each panelist in Round 1 represent his or her best individual interpretation of the patterns without collaboration. The aggregate rating of the whole panel after the final round represents the best collaborative interpretation. Therefore, the experimental hypothesis predicts that, if the panelists benefited from collaboration, then their aggregate ratings in the final round should be more similar to the ratings of the criterion judge than their individual ratings from the first round. In the first round, there obviously was no collaboration, but by the final round, there had been multiple opportunities to consider the ratings and comments of other panelists. Therefore, the main focus of the analyses are the comparison of the criterion rating to the first-round rating of each expert on the panel and to the group's final-round aggregate rating. The comparison of the criterion rating to that of the out-group expert is also included in the data for secondary discussion. In this chapter, the term "rater" refers to the individual experts and the group.

5.2 Kendall Rank-order Correlation Coefficient

Correlations provide a simple measure of relatedness between two variables. Likert scale ratings, such as those used in the present study, fall between ordinal and interval levels of measurement. The more conservative approach for data analysis in these cases is to use an ordinal, non-parametric correlation.

A commonly used non-parametric technique to compare the correlation between two sets of rankings from two subjects is the Kendall rank-order correlation coefficient, often referred to as Kendall's tau [56]. The ratings collected in the experiments were converted into rankings, and the Kendall rank-order correlation coefficients were computed using the group's last round and the expert's first rounds compared to the criterion judge's. The result is shown in Table 8.

At an α level of 0.05, a Kendall's tau in Table 9 that is highlighted in bold indicates rejection of the null hypothesis, meaning that there is significant correlation between the criterion and a rater's rankings. A graphical view of these coefficients, shown in Figure 14, gives a better picture of the tau values relative to each other.

If the hypothesis is correct, that the judging panel collaborating as a whole will make better decisions about the importance of the patterns than any individual panel member, then the correlation between the pattern ratings of the criterion judge and the average ratings of the whole panel should be greater than the correlation between the ratings of the criterion judge and the ratings of any individual panel member. One fundamental issue related to the design of this research is how strong a correlation has to be to conclude that there is good agreement between the panel and the criterion judge. It is not sufficient for the correlation between the panel and the criterion judge simply to be greater than the correlation between the individual judges and the criterion judge; the correlation must be strong enough to indicate substantial agreement. In other words, if the correlation between the whole panel and the criterion judge is weak, it doesn't matter if it

Table 9. Kendall's tau – raters compared to criterion

		G	E1	E2	E3	E4	E5	E6	E7
		r3	r1	r1	r1	r1	r1	r1	r1
April									
N=31	tau	**0.659**	**0.464**	0.235	**0.553**	**0.452**	**0.4**	0.293	
	p	**0**	**0.003**	0.123	**0**	**0.003**	**0.011**	0.056	
May									
N = 22	tau	**0.402**	0.117	0.166	0.288	- 0.117	**0.491**	**0.671**	0.307
	p	**0.017**	0.517	0.369	0.117	0.523	**0.009**	**0**	0.084
June									
N=27	tau	**0.495**	**0.452**	0.214	0.128	0.121	**0.351**	**0.378**	0.305
	p	**0.001**	**0.007**	0.237	0.444	0.48	**0.037**	**0.022**	0.07
July									
N=26	tau	**0.642**	0.215	0.098	**0.414**	**0.416**	**0.606**	**0.632**	**0.415**
	p	**0**	0.222	0.598	**0.017**	**0.015**	**0.001**	**0**	**0.016**
Aug									
N=40	tau	**0.424**	**0.325**	.	**0.551**	- 0.052	**0.603**	**0.313**	0.163
	p	**0.001**	**0.023**	.	**0**	0.708	**0**	**0.023**	0.234

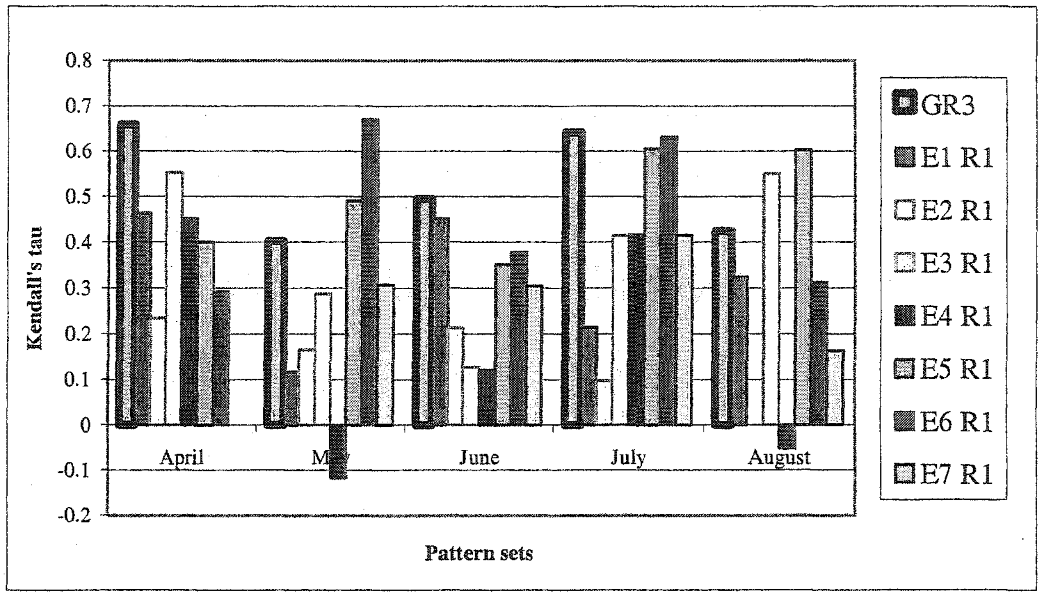


Figure 14. Graphical view of Kendall's tau - rater compared to criterion

is greater than the correlation between individual panel members and the criterion judge, because none of the ratings reflect the judgments of the criterion judge. Therefore the correlation between the whole panel and the criterion judge must be substantial as well as greater than the correlation between individual judges and the criterion judge.

One criterion for judging the strength of correlation is statistical significance [6]. This is a necessary but not sufficient criterion. Correlations can be statistically significant without being particularly informative in terms of prediction and similarity. Groth-Marnat [24] recommends looking at the validity correlations of similar kinds of assessments to determine the level that is considered to be a substantial correlation. In the present research, the judges represent widely varying perspectives, expertise, and clinical experience, and hence varying criteria and priorities in rating the patterns. Therefore, the appropriate comparison for determining the definition of a substantial correlation would be found through assessments of agreement between subjects and a criterion that involve multi-dimensional behavioral judgments. Such comparisons occur most frequently in research on psychological assessment.

Within the framework of psychological assessment, the present research would be classified as a form of criterion validity known as concurrent validity [24]. Criterion validity involves comparing test scores to an external criterion. Concurrent validity is a form of criterion validity in which subject scores and criterion scores are taken at about the same time and compared.

Groth-Marnat [24] points out that as the number of variables that influence a trait increase, the expected validity coefficient will be lowered, due to unpredictable interactions and unaccounted for dimensions of the trait. For example, Neisser et al. [42], in a comprehensive review of the literature on the criterion validity of intelligence tests based on academic performance, concludes that the overall correlation is around 0.5. The criterion validity of intelligence tests is among the highest found for psychological assessments.

The American Psychological Association established the Psychological Assessment Work Group to review issues related to the validity of psychological assessment. In one of their reports [38], they established comparison norms for validity for both psychological and medical assessments through an exhaustive review of assessment research. They reviewed more than 125 meta-analyses including more than 800 samples. They found that the high range for validity coefficients for both medical and psychological assessments was in the 0.5 to 0.6 range, with the bulk of the validity coefficients ranging from 0.15 to 0.3. They concluded that their findings "highlight how challenging it is to consistently achieve uncorrected univariate correlations that are much above .30 [38]." The findings of the current research should be interpreted within this context of criterion validity. Therefore, for the results reported in Table 9, the Kendall rank-order correlation coefficients that are highlighted in bold are considered significant and substantial.

Table 9 shows higher rank-order correlation of the rankings between the group and the criterion than that of the individual for three of the five months. However, none of the individual experts performed better than that standard. In other words, in the two months where an individual outperformed the group, it was a different individual in each month. Therefore, the result is encouraging in that the collective expertise could perform better than any single expert alone.

An alternative to using Kendall's tau as a measure of similarity is to assume that the Likert rating scale yields interval level data that can be analyzed using Pearson product-moment correlations [65]. So, for completeness, Pearson correlation coefficients were computed for pair-wise comparison of the ratings of the group's final round and the individual's first round to those of the criterion judge. The results are shown in Table 10.

The pattern of correlations and significance is essentially the same as that found using Kendall's tau in that the group did better than any single rater for the months of April, June, and July. It is reassuring that the Pearson correlation coefficients show consistent results.

Correlation coefficients measure the relative similarity in ratings, but do not provide an accurate measure of absolute agreement on two measurements. For example, if one judge rated one pattern a 1 and a second pattern a 2, and a second judge rated the first pattern a 4 and the second a 5, then the correlation of their ratings would be perfect even though they really did not agree on the interpretation at all. Bland and Altman [11] make this point and showed that graphical methods provide a better description about the

Table 10. Pearson correlation coefficients – raters compared to criterion

		G	E1	E2	E3	E4	E5	E6	E7
		r3	r1	r1	r1	r1	r1	r1	r1
April N=31	r	**0.718**	**0.459**	0.283	**0.635**	**0.496**	**0.471**	0.299	
	p	**0**	**0.009**	0.122	**0**	**0.005**	**0.007**	0.103	
May N = 22	r	**0.532**	0.099	0.174	0.313	-0.009	**0.495**	**0.736**	0.36
	p	**0.011**	0.661	0.438	0.156	0.968	**0.019**	**0**	0.1
June N=27	r	**0.608**	**0.519**	0.188	0.232	0.022	0.377	**0.5**	0.37
	p	**0.001**	**0.005**	0.347	0.244	0.914	0.053	**0.008**	0.057
July N=26	r	**0.835**	**0.535**	-0.022	**0.449**	**0.51**	**0.759**	**0.768**	**0.495**
	p	**0**	**0.005**	0.914	**0.021**	**0.008**	**0**	**0**	**0.01**
Aug N=40	r	**0.66**	**0.472**	.	**0.62**	-0.105	**0.788**	**0.444**	0.243
	p	**0**	**0.002**	.	**0**	0.519	**0**	**0.004**	0.131

agreement. In the present research, the Kendall's tau does not provide a clear picture of absolute agreement, nor does it provide explanations on where the agreements were. Therefore, other techniques, including a graphical one, were utilized to gain better insight into absolute agreement. The following sections describe the analysis using the other techniques.

5.3 Analysis Using Sum of Mean Difference of Rating

In addition to not providing an index of absolute agreement, another drawback of using Kendall's tau to analyze the data is that there is not a simple way to summarize the overall performance of each panelist and the group for all five months of the Delphi process. A simple way to examine absolute agreement or disagreement is to compute the average difference in ratings between the criterion judge and the panelists as follows:

$$\text{Mean rating difference of expert } E_j \text{ for month k} = \frac{\sum_{i=1}^{n} \text{abs}(CR_i - E_jR_i)}{n}$$

where n = number of patterns in set k, CR_i = criterion rating for the i^{th} pattern, and E_jR_i = expert Ej's rating for the same pattern. The value of k corresponds to the months of April to August. The results for the five months are shown in Figures 15 to 19. In the figures, GR3 means the averages were computed using the ratings of the final round of the group. E_jR1 means the first-round ratings of expert j were used.

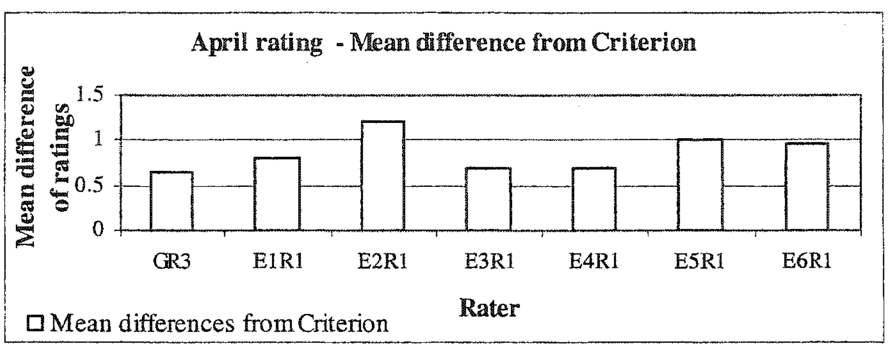


Figure 15. Mean difference from criterion for April 2001

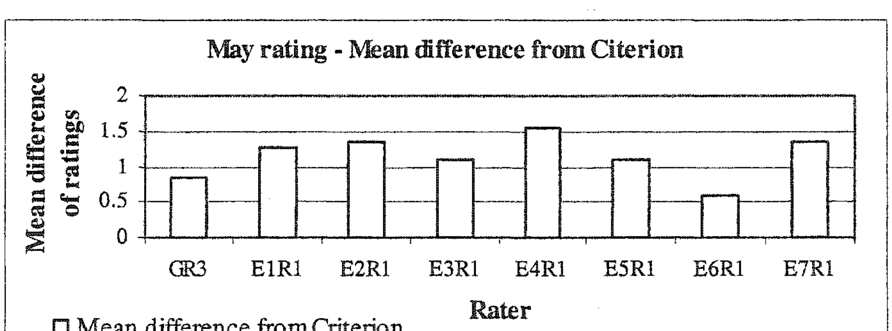


Figure 16. Mean difference from criterion for May 2001

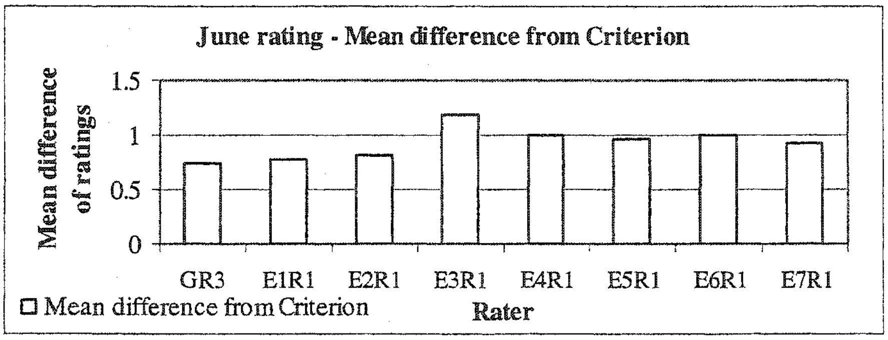


Figure 17. Mean difference from criterion for June 2001

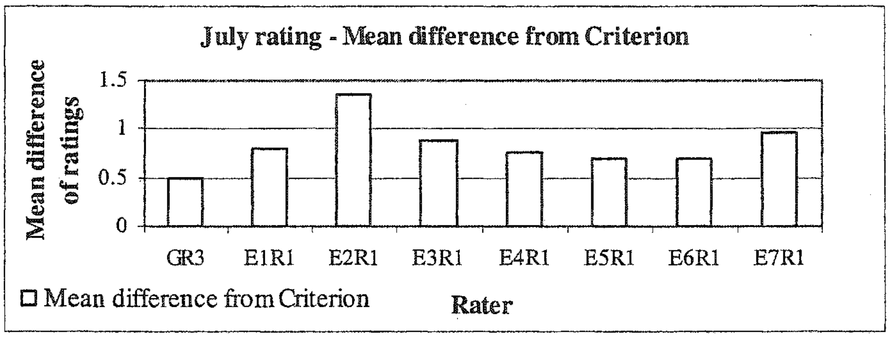


Figure 18. Mean difference from criterion for July 2001

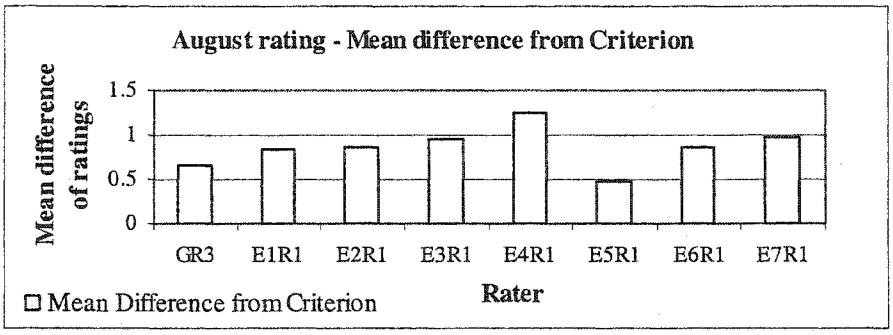


Figure 19. Mean difference from criterion for August 2001

The results from the mean difference analysis is very similar to the Kendall's tau analysis in that the group performed better than every individual in April, June, and July by having the smallest average difference of rating from the Criterion. The benefit of this technique is that the sum of the mean difference over the five months could be calculated easily to provide a summary of the experts' and the group's overall performances. The results shown in Figure 20 indicate that the group has the minimum overall difference in rating from the criterion. That means the group performed better than any single expert overall.

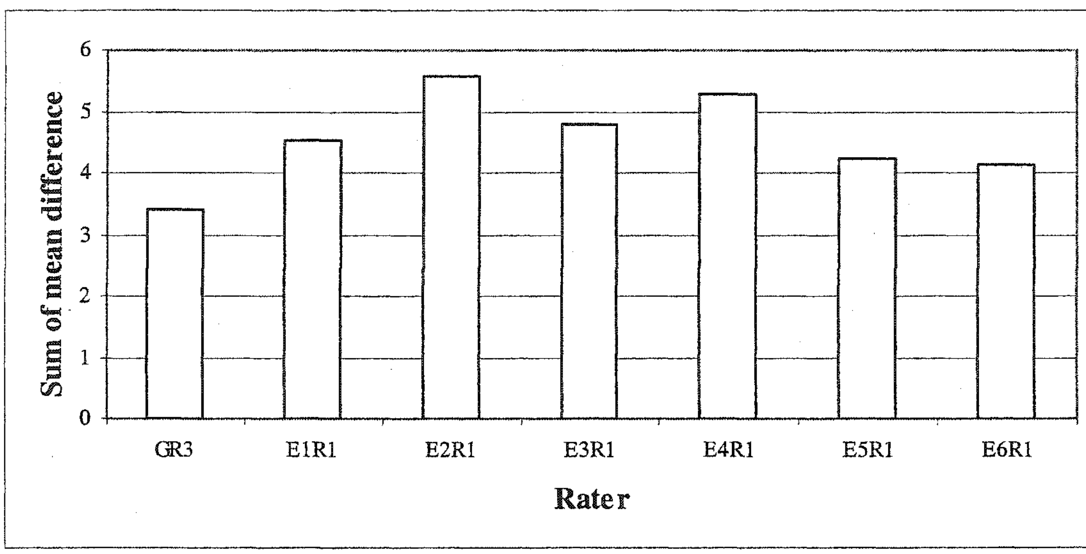


Figure 20. Sum of mean differences for the in-group panel for 5 months

The sum was also computed by including the out-group's rating. Since the out-group member did not join the experiment until the second month, the sum was computed only for the four months that the out-group member participated. The result is shown in Figure 21. The result also showed that the group performed better.

While this analysis reinforces the findings of Kendall's Tau in terms of absolute agreement or disagreement, it still suffers from one shortcoming. This analysis still does not provide insight into exactly where among the ratings the agreement and disagreement occurred.

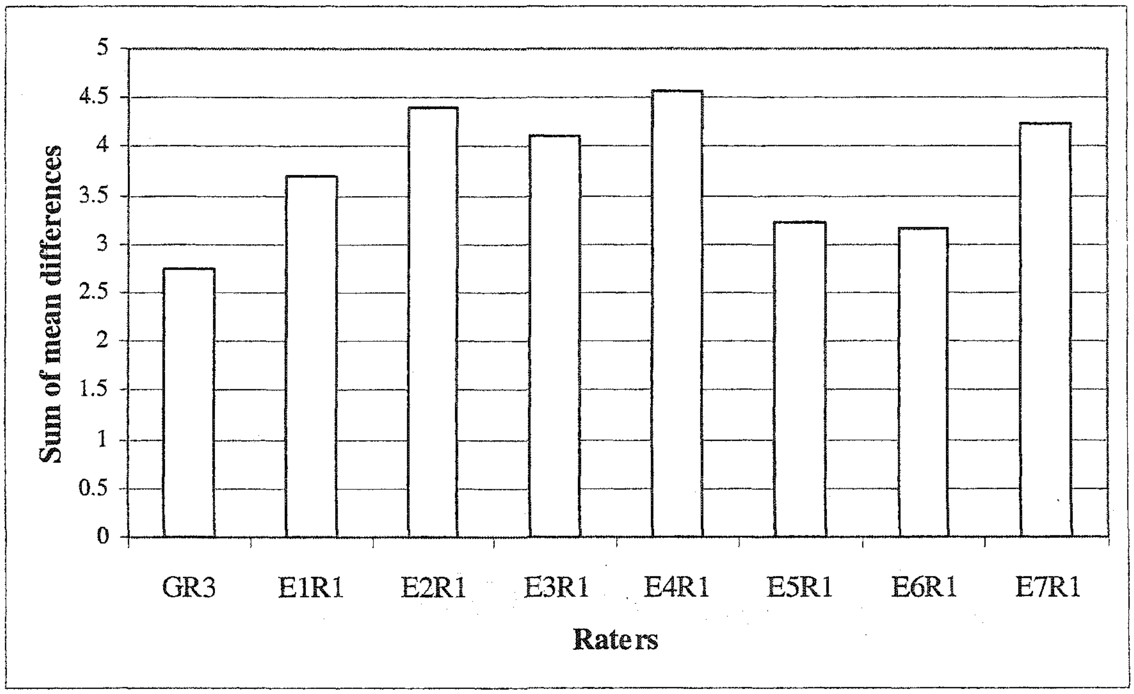


Figure 21. Sum of mean differences of in-group and out-group panel for 4 months

5.4 Analysis Using Graphical Method

The graphical way is most useful in seeing the relationship of the ratings between the group and the criterion judge. Figures 22-26 show the ratings of the group's final round against the criterion rating. The ratings for the patterns of the five months are shown in the figures. The x-axis of each graph are the patterns of the month, represented by the pattern IDs. The y-axis represents the ratings. The patterns are sorted in descending order of importance based on the criterion judge's ratings.

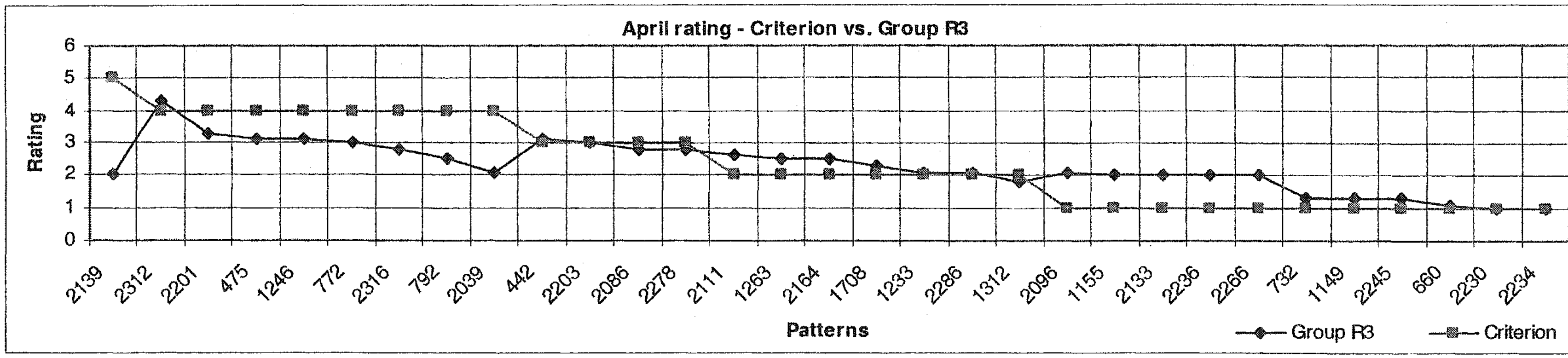


Figure 22. April - criterion rating plotted against Group R3 rating

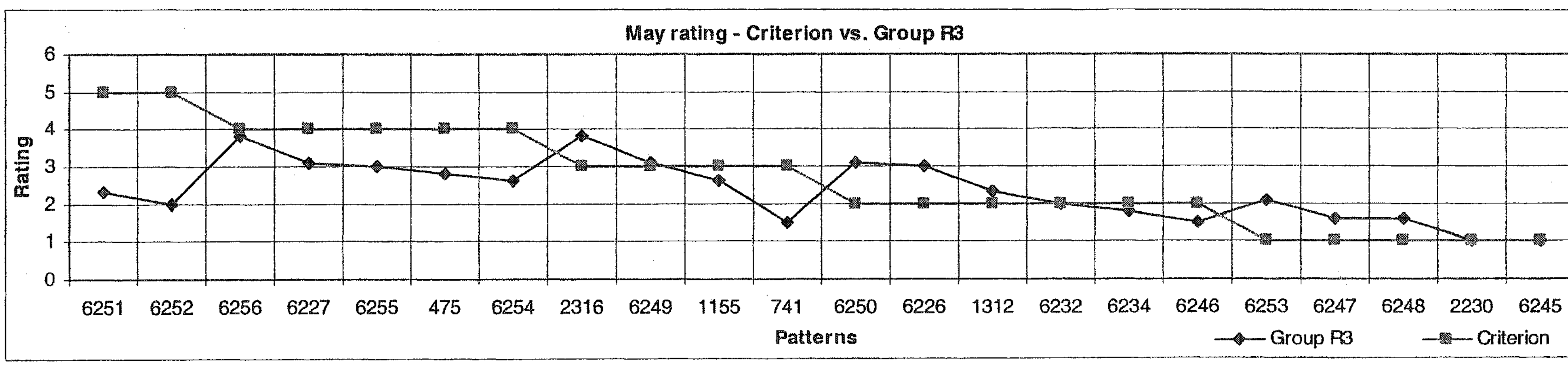


Figure 23. May - criterion rating plotted against Group R3 rating

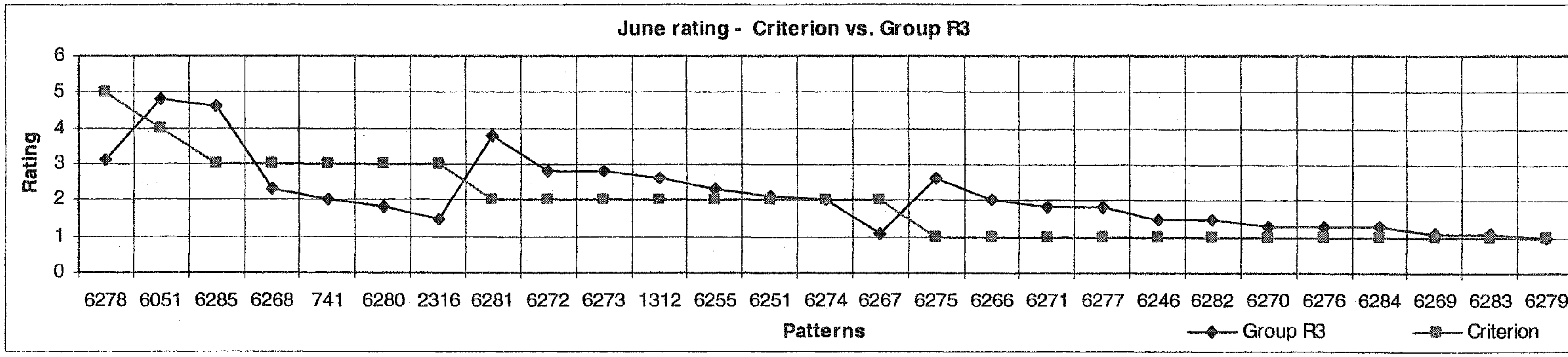


Figure 24. June - criterion rating plotted against Group R3 rating

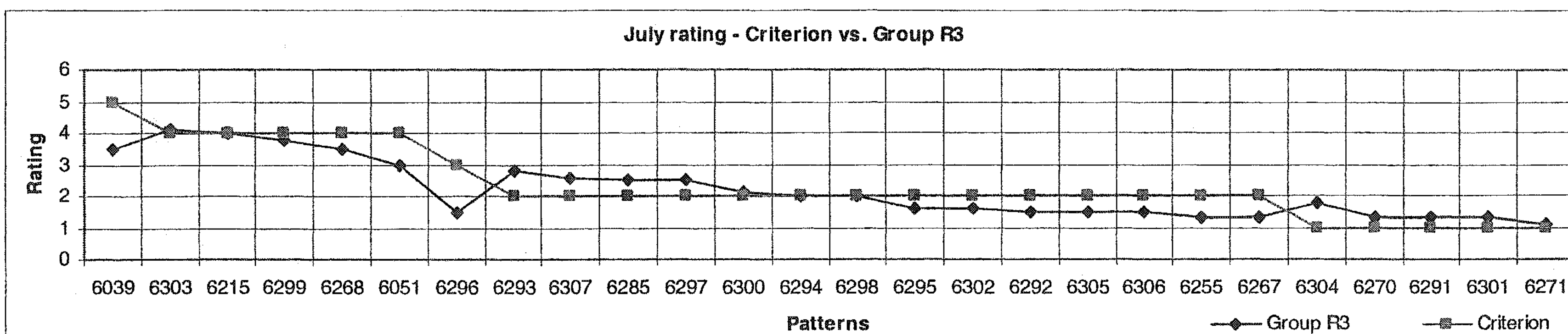


Figure 25. July - criterion rating plotted against Group R3 rating

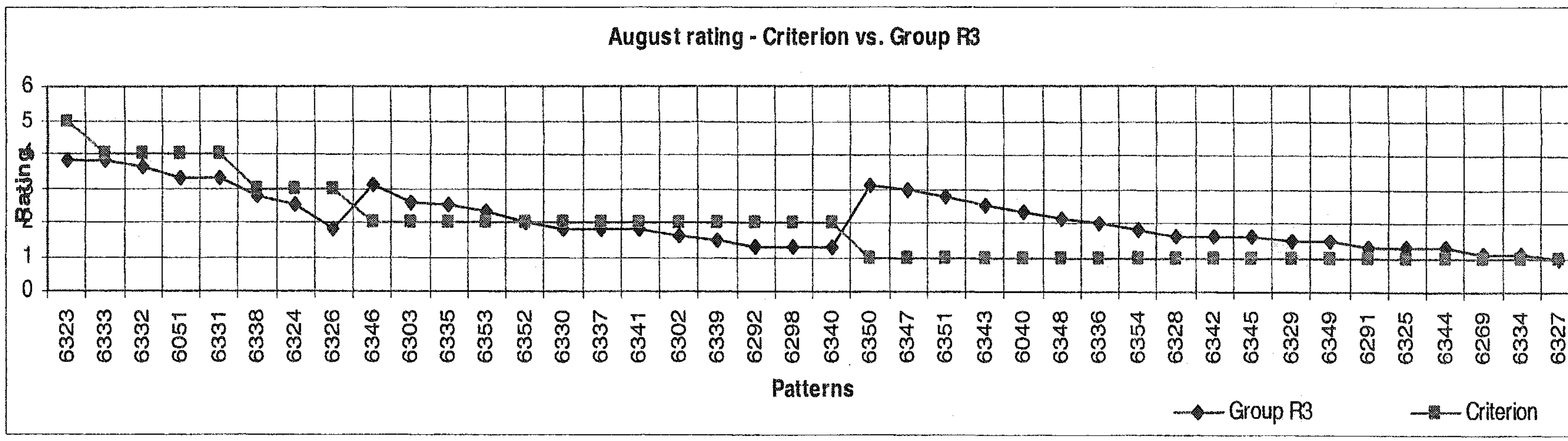


Figure 26. August - criterion rating plotted against Group R3 rating

Both Kendall's tau and the mean differences indicate the presence of disagreement between the group and the criterion. What is critical, and what is displayed here, is the nature of the disagreement. If there is random disagreement among all patterns, then it is unlikely that the disagreements could be easily minimized or resolved. If, however, there is good agreement for most patterns, but striking disagreement for a few patterns, then this finding indicates that such disagreements might be resolved through further collaboration, and the agreement between the group and the criterion judge improved. Furthermore, these key patterns, which lead to marked disagreements, become the focus of further research. What is it about these patterns that leads to such disparate interpretations?

In general, the following were observed in the five months of data using the graphs:

1. There is a high level of agreement for unimportant patterns, i.e. patterns with ratings 2 or below. The number of unimportant patterns in a month ranged from 36% to 50%, representing an average of 45% of the patterns in the experiment.
2. There is reasonable agreement as to the important patterns, i.e. patterns with ratings above 3.
3. There were a few key patterns in which the difference between the group and the criterion was greater than 2 and indicated that the criterion considered the pattern important but the group did not.

The most striking observation from these five graphs is observation 1, which indicates there was good agreement on which patterns were not important. This is an interesting finding because it suggests an opportunity for refinement of the data-mining algorithm to reduce uninteresting patterns. However, a useful metric is still needed to examine the hypothesis. The next section described an adaptation of the recall and precision metric from the Information Retrieval field to examine the agreement on the important patterns.

5.5 Analysis of Important Patterns Using Hit Rate and Precision

The goal of the collaboration is to identify the important patterns in a set for the purpose of further epidemiologic investigation. Therefore, it is important to gain a better understanding of how well the group, versus the individual, performs in terms of identifying the patterns that are considered important by the criterion judge. In addition, it is also useful to know the level of false positives identified by each group because it indicates how precise the judgment is. A false positive, or false alarm, is a pattern that the group rates as important, but the criterion judge does not. None of the three techniques discussed above in Sections 5.2 to 5.4 provide such insight. The graphical method helps to reveal the trend of agreement and disagreement, but it does not provide a measure to compare the individuals to the group.

In order to gain a better measure of the degree of agreement between the two sets of ratings on the important patterns, the metrics of recall[9] and precision[10] from the field of Information Retrieval [52] are utilized here. For this analysis, an important pattern is defined as one that the criterion judge has rated above 3. An unimportant pattern is one

---

[9] Recall: A standard measure of IR performance, recall is defined as the number of relevant documents retrieved divided by the total number of relevant documents in the collection.

[10] Precision: A standard measure of IR performance, precision is defined as the number of relevant documents retrieved divided by the total number of documents retrieved.

that the criterion judge has rated a 3 or lower. The two corresponding metrics defined here are hit rate and precision.

In Figure 27, the left-hand clear circle represents the set of patterns that the criterion judge considered important in a given month. The right-hand dotted circle represents the patterns that were judged important by an expert in the first round, or by the group after the final round for that same month. The overlapping area of the two circles, highlighted with diagonal stripes, represents the set of patterns that are rated important both by the criterion judge and an expert or the group. These patterns are called hits, and they indicate agreement between the criterion judge and the panel that a pattern is important (rated above 3).

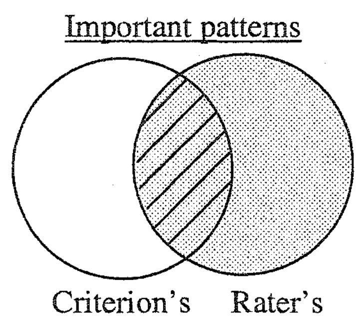


Figure 27: Sets of overlapping important patterns from criterion and a rater

Hit rate measures the completeness (accuracy) of the selection of important patterns. Precision measures the purity of the selection. High precision indicates that the ratings are not cluttered with false alarms. For this domain, it is important to have high hit rate, and high precision is desirable as well. The hit rate and the precision are computed as follow:

$$\text{Hit rate} = \frac{|\,\text{Criterion} \cap \text{Rater}\,|}{|\,\text{Criterion}\,|}$$

$$\text{Precision} = \frac{|\,\text{Criterion} \cap \text{Rater}\,|}{|\,\text{Rater}\,|}$$

| Criterion ∩ Rater | is the number of patterns in the set of hits of a rater. A rater can be an expert or the group after the final round. | Criterion | is the number of patterns in the set of important patterns of the criterion judge. | Rater | is the number of patterns in the set of patterns rated important by the rater. Using these two measures, the hit rates and precisions using each expert's first-round rating and the group's final-round ratings were computed for each month's pattern set. The average of the five months for each measure was also computed. Figure 28 shows the average hit rate for the five months, using ratings from the group's final round and the individual's first round. Figure 29 shows the average precision of those five months.

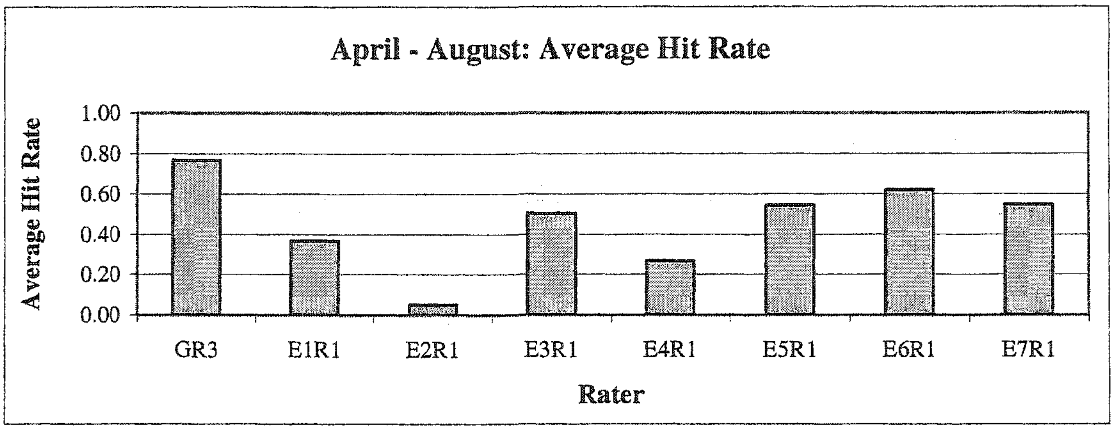


Figure 28. Average hit rate for the five months

The result in Figure 28 show that the group out performed everyone in hit rate. Figure 29 show that the group is second in precision. A rater in total agreement with the criterion judge would have maximum hit rate and precision. When total agreement is not achieved, however, a small number of false alarms is far more acceptable than misses in this domain. Therefore, these data again show that the group performed better than any single individual in identifying the important patterns.

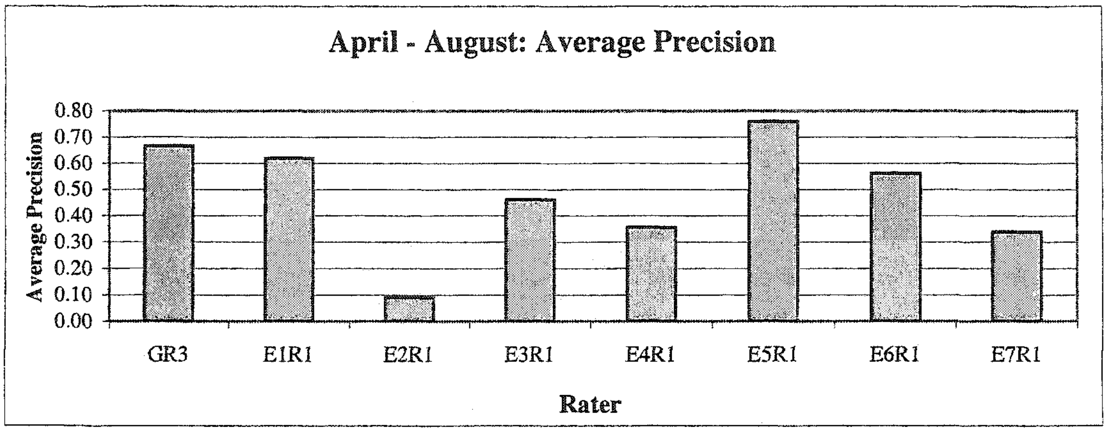


Figure 29. Average precision for the five months

5.6 Observations from Experts' Reasons

Further observations regarding agreements, disagreements, and influence on individuals through collaboration were obtained by reviewing the rating changes between rounds and the elaborated reasons submitted by the experts. Examples are given here to illustrate the observations. The example figures are snapshots of the researcher's user interface screen. The histograms are the rating distribution of round 1, round 2, and round 3. The rounds are listed from left to right in the figures. The bold numbers in the left-hand column of each round are the rating scale of 1 – 5. The small numbers at the right of a histogram bar are the expert identification numbers E1 – E6, except the "E"s are not displayed on the screen. The Score is the group average rating. The Group Rank is the ranking of the pattern derived from the group's average rating.

1. Strong agreements starting at round 1.

   There were patterns on which the panel had good agreements from the start. Figure 30 shows a pattern where every expert considered it unimportant in round 1. The group did not change its position after collaboration. This shows that the pattern is trivial. The group's unanimous agreement starting at Round 1 indicates that this pattern is a good candidate to be eliminated by the DMSS to reduce the amount of trivial patterns generated. Figure 31 shows a pattern

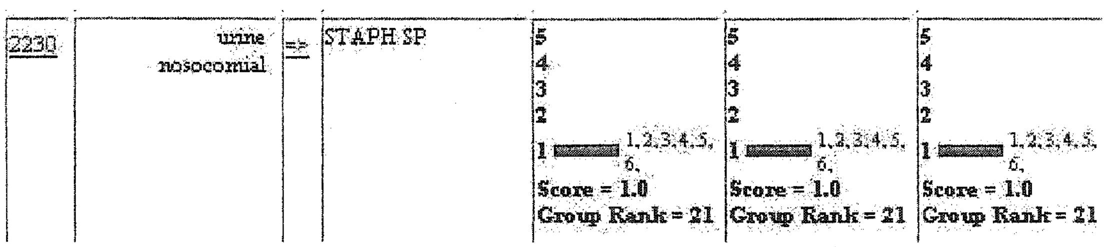


Figure 30. Unimportant pattern with unanimous agreement of the panel

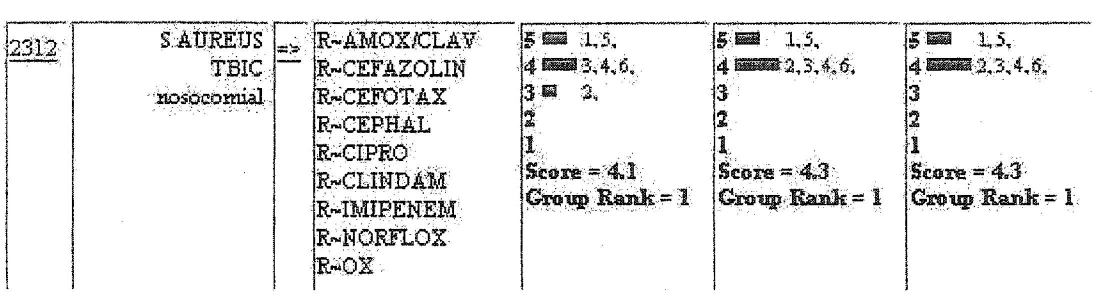


Figure 31. Pattern considered important even before collaboration

that the panel in general agreed to be important starting in Round 1. The slight change of position by E2 in round 2 only serves to reinforce the group's opinion.

2. Disagreement remain after collaboration

   Figure 32 is an example that shows that CIS does not force consensus but captures the differences in opinion. In round 1, the group's opinion on the importance of the pattern was quite diverse. Starting in round 2, with the exception of E4, most of the group's ratings gravitated to a 2 and 3 rating. Such rating distribution indicated that most of the group considered the pattern not very important. One of the panelists, E6, submitted the following reason:

   > *Check for Cystic Fibrosis (CF) patients.*

   In UABH, most CF patients are admitted to J3E. However, non-CF patients could also be admitted to J3E. CF patients are prone to infections by drug resistant organisms. E6's comment showed that the group might have assumed that the patients in J3E were CF patients. However, E4 remained adamant that it was an important pattern and offered the following reason:

   > *All of these except the two from 5/2001 were resistant/inter. to drugs that we routinely use to treat this infection ie Cipro, Ceftaz, Pip. I would be limited to using drugs that are perhaps more toxic for the patients and more expensive for the hospital. Since Cipro is the only oral drug to treat this organism, outpatient therapy would be difficult because I would be forced to use IV treatment should they require additional therapy after D/C. Given that these patients were on J3E, it is possible that they were CF patients which pose quite a challenge to the caregiver with regard to abx therapy. It is important they receive abx with good coverage of this organism and be dosed in a way that would limit or slow the progression to more resistant strains.*

This comment indicated that E4's concern was from the treatment point of view. This expert considered that it was important to review the patients' treatment immediately in order to prevent the organism to evolve into more resistant stains, which would make treatment more difficult. This example illustrates the importance of capturing the differences in opinion and not forcing a consensus in an expert collaboration process. The reason submitted by E6 also shows that knowledge about the hospital operation, such as the location where certain types of patients are admitted to, is important to decision-making.

ID	Surveil	=>	Outcome	Round 1	Round 2	Round 3
475	JBE resp nosocomial	=>	P AERUGINOSA	5 4, 4 2, 3 3, 2 1,6, 1 5, **Score = 2.8** **Group Rank = 4**	5 4, 4 3 2,3, 2 1,5,6, 1 **Score = 2.8** **Group Rank = 8**	5 4, 4 3 2,3, 2 1,5,6, 1 **Score = 2.8** **Group Rank = 8**

Figure 32. Disagreement remains after collaboration

3. Agreement reached after collaboration.

Figure 33 shows an example where there were differences in opinion among the experts, as indicated by the difference in their ratings in round 1. E5 has submitted a reason in round 1 to support his or her importance rating of 5. The reason was:

> *An outbreak due to transmission seems likely because of unusual pattern, no. of outcome isolates too high, and cluster by time and location.*

The location, R3, specified in the pattern is a rehabilitation ward. Such kind of infection is considered unusual for this location and is an indication of transmission. E2, E3, and E4 initially did not consider the pattern important. According to the roles of these experts, they do not provide direct care to the patients in R3 and hence they might not have paid attention to the events at that location. The ratings of E1, E5, and E6, together with E5's reasons, drew E2, E3, and E4's attention to this pattern. They eventually changed their rating to concur with E1, 5, and 6. This change of position is an indication of influence on individuals by other's experience and knowledge on issues that the individuals may not be familiar with or particularly confident in. This showed that collaboration facilitated knowledge sharing.

6051	R3 nosocomial	=>	S AUREUS R-AMOX/CLAV R-CEFAZOLIN R-CEFOTAX R-CEPHAL R-CIPRO R-CLINDAM R-NORFLOX R-OX	5 5,6, 4 1, 3 2 1 2,3,4, **Score = 2.8** **Group Rank = 3**	5 2,3,5,6, 4 1, 3 2 1 4, **Score = 4.1** **Group Rank = 2**	5 2,3,4,5,6, 4 1, 3 2 1 **Score = 4.8** **Group Rank = 1**

Figure 33. Pattern showing agreement reached after collaboration

4. Overall comments from experts review difference in view points.

The following non-pattern specific comments submitted by the experts reveal opposing views in regard to the importance of patterns in general:

E5: I believe the Group attaches (1) more significance than I do to small numbers of gram-negative organisms that may or may not be in special populations (e.g., cystic fibrosis) or related by transmission, and (2) less *importance to substantial numbers of highly-transmissible Staph. aureus in a single unit.*

E2: This is perhaps because our patients with MRSA are generally already on strict contact isolation to reduce spread of infection, and while MRSA is a significant infection, it is perhaps perceived to be less virulent than most of the gram negative organisms encountered (the more resistant gram negatives generally require 2 antibiotics, whereas we can presently still treat MRSA with 1)

E2's comment was a response to E5's. E2's comment shows that the expert approached pattern interpretation from a treatment point of view, whereas E5 came from the viewpoint of transmission prevention. E2 is a critical care physician, and E5 is an infectious disease specialist. The viewpoints each took were consistent with their roles.

5.7 Summary of Data Analyses

The Kendall's rank-order correlation coefficient (tau) showed substantial relative agreement between the group's ranking and the criterion ranking. That is, patterns that were rated relatively high by the panel as a whole were, in general, also rated high by the criterion judge, and vice versa. Furthermore, the correlations between the aggregate group ratings and the criterion were more similar in general than the individual rankings. While this finding supports the hypothesis, correlations are not an indicator of absolute agreement and do not provide insight about the characteristics of the agreement.

The sum of mean difference analysis is simple to compute and provides an easy way to summarize the disagreements over all of the months. The monthly result from this analysis is consistent with the result from the Kendall's tau. The summary over the five months helps to show that the overall performance of the group is better than any of the individual. This technique also suffers a lack of specificity to describe the characteristics of the agreements or disagreements.

The graphical method of plotting the ratings of each pattern provides insight into the characteristics of the agreement. The graphs showed good agreement on the unimportant patterns and reasonable agreement on the important ones. It is the ones in the middle rating range that the ratings varied and hence likely the cause of the moderate correlation coefficients. A drawback of the graphical method used here is that the graphs become very busy visually when the ratings of all of the raters and the criterion are plotted on the same graph. Also, it becomes difficult if the number of patterns is large. Since the goal of the experts' task was to prioritize the patterns so that the important ones could be identified for infection investigation, it is desirable to be able to measure the extent of agreement on the patterns that the criterion considered important. Unfortunately, the graphs do not provide a simple way to measure how good the agreements were for the important patterns, i.e. those with ratings greater than 3.

The hit rate and precision analysis provide metrics to measure how well each rater performed in identifying the important patterns. The result was encouraging. However, the metric lacks a standard to determine at what level the rater's performance is considered good. Nonetheless, the method provides a means to compare the relative performance between the raters.

Kendall's tau was computed using the rankings derived from the ratings. The other three techniques analyze the ratings directly. Unlike the Kendall rank-order correlation coefficient, the other three techniques lacked a formal mechanism for testing statistical significance. However, the four techniques augment each other in providing insight into the agreement. Together they show that the group has a more consistent pattern of both relative and absolute agreement with the criterion than any individual. In addition, the group in general performed better than the individuals in identifying the important patterns.

The four empirical analyses were augmented by reviewing the reasons submitted by the experts. The reasons reveal the differences in opinion that reflect the role of the experts in the organization. The reasons submitted by the experts were also a medium of knowledge sharing among the panelists.

5.8 Mid-study Survey

A mid-study survey was conducted after two months of Delphi. All of the surveys were conducted by face-to-face interviews except two. Those two were conducted over the phone. The goals of the survey were as follows:

1. To find out what DMSS data the user relied upon to make decisions.
2. To find out what feedback data the user relied upon to make decisions.
3. To find out if there were unreported problems or difficulties using the CIS.
4. Provide an opportunity to do additional training if requested.
5. To listen to the user think out loud in interpreting the patterns.

The experts' responses to the survey questions are summarized in Appendix D. In general, the group used the following information the most for making decisions, starting from Round 1:

1. Pattern description.
2. Surveillance and outcome isolate counts.

3. Location.

In general, the responses showed that the group used the following feedback the most for decision making from the second round onward:

1. Group score distribution
2. Elaborated reasons

5.9 Post-study Survey

A post-study survey was conducted upon the completion of the fifth month's pattern interpretation. The purpose of the survey was to collect the users' opinions in regard to the usability of the CIS and the Delphi process for decision-making. The survey was conducted via the web. The survey processing used the CIS group processing software with the maximum rounds for the Delphi process set to 1. The survey responses are summarized in Appendix E. The following are some CIS usability general findings:

1. Easy to use and understand (4.4/5, 5=strongly agree)
2. Response time (4.4/5, 5=strongly agree)
3. Usefulness of other's rating and reasons (4.1/5, 5=strongly agree)
4. Difficulty of making decision using the information (2.9/5, 5=very easy).

The survey result showed that the CIS was considered easy to use in general and therefore was not a hindrance to the collaboration and decision-making process. However, the group thought that it was not easy to decide on the importance rating using the DMSS pattern information, as indicated in finding 4 above. This really underlines the fact that interpretation of the data-mining result is not as easy as generally assumed by KDD researchers.

Some users commented that the system provided a knowledge-sharing environment. As expressed by one expert, it "gives one the opportunity to initially make a decision independently, but subsequently re-examine it in light of other's input. Good for learning." Another user advised, "Don't stop now."

5.10 Other Observations

A number of CIS usage and user behaviors could be observed using the results and the usage logs. They are summarized in the following sections.

5.10.1 Changes at Each Round and Group Influences

Group influence on individuals was observed based on the percentage of importance rating changes. The observed changes are shown in Figure 34 for each pattern set. The data showed that most changes of opinion occurred in the second round, which was consistent with our prediction.

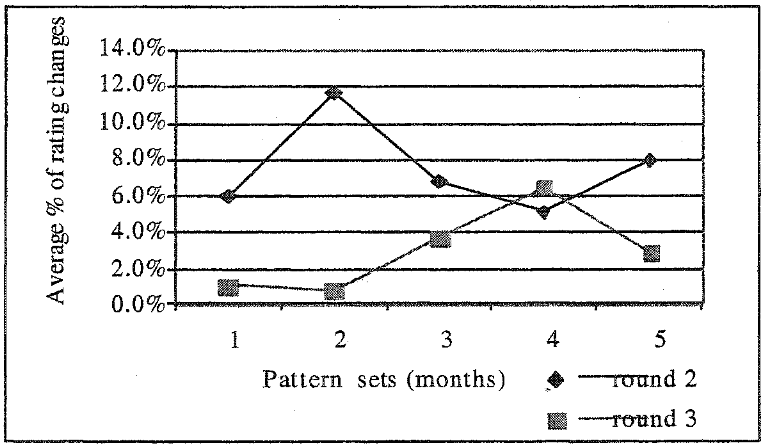


Figure 34. Average of rating changes

5.10.2 Use of Defaults

On the average, 33% of the patterns remained to have the default important rating of 1. This is another indicator for uninteresting patterns. These data support the observations discussed in the graphical analysis in section 5.4 that there is good opportunity for the refinement of the data-mining algorithm to reduce the amount of uninteresting patterns. The reduction of uninteresting patterns will make the interpretation task more achievable and would likely improve the quality of the interpretations both by the group and by the individuals.

5.10.3 Time to Decision

The system recorded the time when a user logged into the system and the time when a user submits his or her ratings by pressing the Submit or Re-submit button. For each login that led to a submit of ratings, an estimate of "time to decision" is computed using the differences of these two times. The average time to decision is the average of the time to decision of the panelists. These averages are conservative because the time incurred between login and submit might include idle time when the user was interrupted by other tasks on their job. The average time to decision for each month and for each round is shown in Figure 35.

The graph shows that the panelists took the most time in the first round to review the patterns and to arrive at a decision. The times spent in the subsequent rounds were substantially lower. This usage of time is consistent with prediction. Only the amount of time spent in the first round was proportional to the number of patterns in the month. The August pattern set has the highest number of patterns (40). The April data set has 31

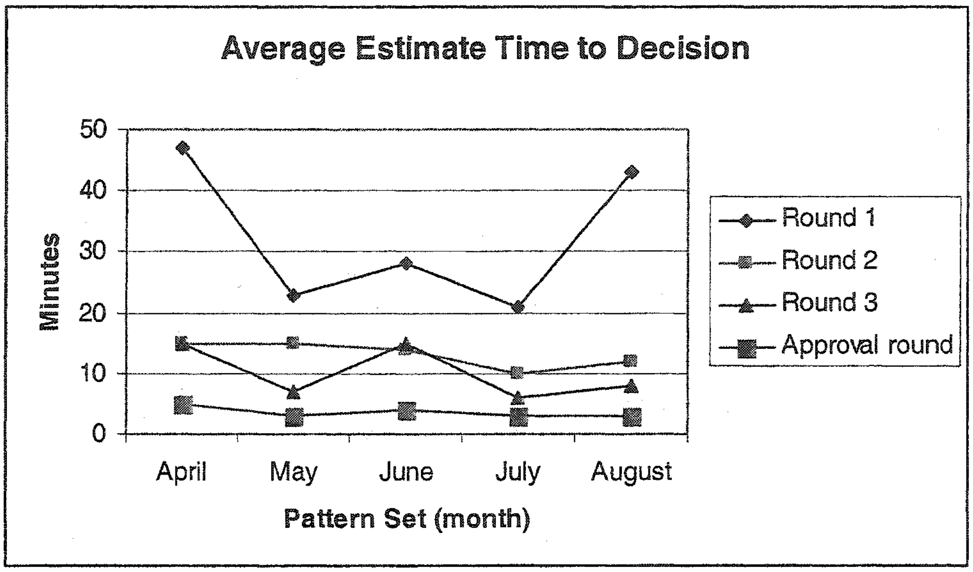


Figure 35. Average estimate of time to decision

patterns, which is the second highest number of patterns in a month. Since April was the first time the panel performed the interpretation, the extra time used was expected.

5.10.4 Usage Patterns

Experts have used the system at various hours of the day and day of the week. The experts reported that they have used it from offices, hospital floors, and in hotel rooms while traveling. Therefore, CIS has achieved the objective of supporting interpretation of geographically and temporally dispersed experts.

The CIS Delphi process was repeated using five months of retrospective data. Our original estimate was that it would take a month for the panel to complete one set of pattern. It turned out that only the first pattern set took a month to complete and that was conducted in a December over the holidays. The subsequent pattern sets took on the average of two and a half weeks to complete. Since we were using retrospective data, therefore, we were able to start a new set as soon as one set was completed. Therefore, we were able to run five Delphi studies within four months. Considering that the panelists were volunteers, the ability to run five Delphi studies in four months is very encouraging. If the participation were the panelists' regular duty, the time for each Delphi study could decrease. Hence, allowing CIS to provide timely interpretation as patterns are available.

5.10.5 Social Loafing

One unexpected finding in the results was the occurrence of social loafing by one member of the panel. Social loafing occurs in groups when one member of the group fails to put forth a strong individual effort, allowing instead for the group to do the work. Potential social loafing was observed in E2's ratings after the second pattern set. This panelist would submit his or her first round without changing the importance rating of any patterns and then selected ratings in the second round. Social loafing is a well-documented behavior in groups [33]. Accountability measures, such as peer rating of one's input, could be added to the CIS to deter social loafing.

CHAPTER 6

DISCUSSION

The purpose of this study was to test the hypothesis that a panel of judges with diverse expertise in the problem domain, yet all qualified to interpret the data-mining results, would yield more effective interpretations when collaborating as a group than any of the judges acting alone. This hypothesis is a direct challenge to the commonly held assumption in the KDD research community that interpretation of discovered patterns is intuitive, trivial, or can be done adequately by the KDD researchers themselves or by a single domain expert. Furthermore, the experiment involved the development, utilization, and assessment of the CIS, which facilitated collaboration among judges who were distributed both geographically and temporally, thus eliminating the need for face-to-face meetings for the purpose of judging data-mining patterns.

The experiment to test the hypothesis was conducted using the CIS prototype for the domain of hospital infection surveillance. The data-mining patterns used in the research were generated by the DMSS using retrospective data from the UABH clinical LIS. The CIS adapted the Delphi Method to provide a web-based, asynchronous, iterative, and structured environment to collect, compile, and archive the expert opinions on the importance of the DMSS patterns. The hypothesis was tested by comparing the group's importance ratings from the last Delphi round and from the individual expert's first round to a set of criterion ratings. The criterion ratings for the pattern set were established using a criterion judge who was supplied with additional data collected by extensive patient chart reviews.

The analyses of the results using the four analytical techniques described in Sections 5.2 to 5.5 found that the group had a more consistent pattern of both relative and absolute agreement with the criterion than any individual. In addition, the group in general performed better than the individuals in identifying the important patterns, and there was good agreement on patterns that were unimportant. These findings in general support the hypothesis.

The results further showed that the CIS is an effective vehicle to facilitate the collaboration process. The post-study survey showed that the CIS user interface is generally easy to use and the response times were good. Therefore, the CIS is not a hindrance in the interpretation and the collaboration process. As a matter of fact, the collaboration of the panel that was distributed both geographically and temporally would not have been possible without the support of the CIS. In the mid-study survey, one of the experts observed that to do such iterative interpretation without the CIS would have been impossible and a nightmare.

The findings of this research offer a number of implications for the field of KDD and are described in the following sections.

1. Collaborative interpreting enhances data-mining pattern interpretation.

The research findings indicate that data-mining output interpretation is more complex than generally assumed by most KDD researchers. Most KDD researchers have either considered their data-mining results to be so intuitive that the researchers themselves were sufficiently knowledgeable to perform the evaluation, or that a single domain expert was sufficient. This is evidenced in research reported in the conference proceedings of the International Conference on Knowledge Discovery and Data Mining for the past eight years [4, 10, 20, 21, 27, 28, 49, 58], and others [2, 19, 34, 67, 69].

The present findings support Uthurusamy's [66] view that KDD should be a collaborative discovery process by a widely distributed team of experts and analysts. This research obtained empirical results indicating that the interpretation of data-mining output can be improved by allowing a group of distributed multi-discipline domain experts to collaborate on decisions about the discovered patterns. Stedman [62] described the difficulty that organizations who use data mining have in retaining a team of data-mining-knowledgeable domain experts to review data-mining results. With a system like CIS, an organization would not need to retain a team of experts just for the purpose of sorting out the implications of the data-mining output. Instead, experts could collaborate in interpretation as part of their job.

2. CIS facilitates continuous knowledge acquisition.

In the collaboration process, knowledge is shared. Such knowledge sharing enhances the individual's knowledge base. Turoff [64] suggested the use of the Delphi Method for continuous knowledge acquisition from multiple experts. The results from collaboration form the collective knowledge of the experts that can serve as a knowledge base for the organization. Organizations are interested in creating organizational memories to retain their intellectual properties {Huynh, 1994 #103;Ackerman, 1994 #85}. The CIS can serve as a mechanism for continuous knowledge acquisition for such a knowledge base.

3. A collaborative group may be able to serve as the gold standard of interpretation.

In the present study, the establishment of the criterion ratings by a single expert using additional data from extensive patient chart review was possible because retrospective data were used and research funding was available. If an integrated system of DMSS and CIS were used on a routine basis for infection surveillance on prospective data, extensive chart review would not be feasible or practical. Therefore, the notion of a gold standard judge would not exist. The experimental results showed that the collective interpretations of a group of diverse domain experts have the potential to perform better than a single expert consistently when complete data concerning the validity of the data mining results are not available. Therefore, the results from collaboration is a good substitute for a gold standard judge with complete information. The implication of this finding is that CI should be part of the KDD toolset. However, even though the criterion judge used in the experiment was more experienced and knowledgeable in infection control, and was provided with the additional data to assist his judging, nonetheless his ratings were still subjective judgments. Therefore, the ability of a panel of experts to identifying important patterns should be further tested by conducting the experiment using prospective data and performing a complete investigation as the events happen.

4. Web services can be used for Group Support Server implementation.

Implication 3 above supports the inclusion of CIS capabilities in a comprehensive KDD toolset. The use of the web-based CIS prototype for the experiment demonstrated that the web is an appropriate channel to support the collaboration. Hence, future CIS development should continue to explore the full potential of the web. An enterprise level development and deployment platform like J2EE was demonstrated to be suitable for the implementation of CIS. The more recent enterprise development platforms -- such as .NET from Microsoft and the latest J2EE, which includes web service support -- should be explored, especially for the development of the domain independent Group Support Server of the CIS. Since the Group Support Server implements the Delphi process and is domain independent, implementing it as a web service could make it more accessible to other applications that need a Delphi-type collaboration mechanism.

5. CIS identifies patterns of interest from disagreements.

The CIS system implemented an adaptation of the Delphi method that facilitated the panel of experts to critically examine the data-mining patterns and collectively determine the epidemiologic importance of these patterns. The Delphi Method does not force consensus but allows the CIS to capture both agreements and differences in opinions. The patterns with difference in opinions should be reviewed with special interest [53]. To the users of the CIS interpretations, such as the hospital infection control staff, the differences in opinions may reveal important factors that should be considered in determining the epidemiologic investigation priorities. In addition, the reasoning that supports the opposing opinions should be examined carefully to gain insight into the decision factors and the potential relationship to the judges' background, such as their roles in the organization, education, and experience.

6. CIS elicits complex decision factors.

A review of the panel's individual importance rating and elaborated reasons, reported in section 5.6, revealed that the experts' judgments tend to follow two viewpoints: prevention or treatment. The treatment viewpoint considers an increased trend of infection or drug-resistance not as important if they know there are other drugs that can be used to treat the disease. The prevention viewpoint considers such an increase as an important event because they prefer the infection not to have occurred. It turned out that viewpoints were consistent with the role of the expert in the hospital. The critical care physicians and the pharmacist were more treatment oriented, whereas the infectious diseases physicians and the microbiologist were more prevention oriented. This means that if only one expert were used to interpret the data-mining result, his or her role and experience would certainly bias the interpretation. In addition, from the mid-study survey, the experts indicated that the location information was useful in their decisions. These observations show that there are various factors influencing the interpretations. There may be additional factors influencing interpretation that the data analyses did not identify or the CIS prototype was not able to capture. The implication is that the initial CIS training process could be modified to elicit some of these factors to be shared and agreed on by the panel before the interpretation task. The patterns with opposing opinions in this study, as described in implication 5 above, could be used in the CIS training of future research for the elicitation of different decision factors.

7. CIS facilitates knowledge sharing.

Both the mid-study survey and the post-study survey indicated that some experts felt that the other experts' input was useful to their decision-making. This is also evidenced in the experts' changes in ratings after the first round. The result showed that even the expert who did not think the group's opinion was useful agreed with the criterion rating more after he or she made changes in the second rounds. This suggests that even the most confident expert could still benefit from the group's collective knowledge.

In this experiment, if a pattern present in a particular month had been generated in a prior month, the CIS did not display the group's judgment from the prior month, in order to minimize influence on the panel's judgment. Future experiments should consider displaying the prior month's judgment as the initial feedback so that the utility of the accumulative knowledge could be evaluated.

8. CIS results serve as feedback for data-mining refinement.

The results reported in the graphical analysis of the data indicated high agreement for unimportant patterns. This is a significant finding because it shows the opportunity for refining the data-mining system to reduce the amount of uninteresting patterns. The collaborative interpretations collected by the CIS could be used to formulate the refinement algorithm. The interpretations could serve as prior knowledge in a feedback mechanism from interpretation to data mining shown in Figure 36.

9. Social loafing is a potential problem.

The CIS prototype implemented a democratic Delphi process that did not intervene in the level of participation by the experts. As a result, potential social loafing occurred, as evidenced in E2's performance from June to August. This shows accountability measures should be considered in CIS to prevent social loafing. A potential method of implementing accountability is for each expert to assess the opinions of the other experts in the form of confidence levels. Such a kind of collective scoring for assessing expert opinions is described in Ayyub [7]. A possible drawback for such a method is that it is time-consuming. The additional time required might deter experts from participating. Other methods for accountability should be explored to prevent unproductive participation due to social loafing.

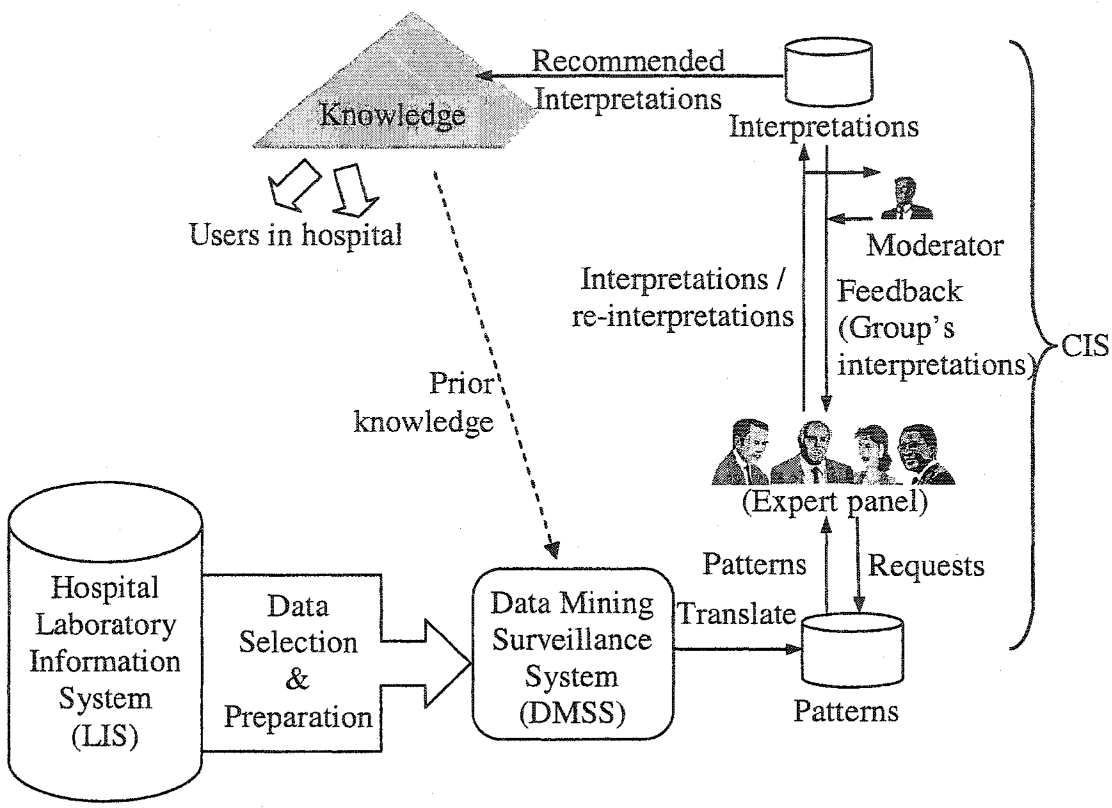


Figure 36. A surveillance framework using DMSS/CIS with feedback to DMSS

6.1 Conclusion

These findings are encouraging in that CIS can 1) utilize collective expertise to enhance the interpretation of KDD results; 2) enrich the knowledge base of experts through collaboration with other experts; 3) provide a mechanism for continuous knowledge acquisition from multiple experts in an organization; and 4) potentially help refine data-mining algorithms to reduce generation of trivial patterns. The use of the DMSS with CIS presents a new paradigm for epidemiological surveillance that is more comprehensive than the ideal computer-based public health surveillance system that is described by Dean et. al. [16].

This research showed the benefit of CI for infection surveillance by establishing subjective criterion ratings for the evaluation of the expert panel's judgments. Further experimentation using prospective data with resources for full epidemiologic investigation is the next logical step to better evaluate how a panel of experts performs in interpreting the data-mining patterns.

As pointed out in the implications above, the CIS prototype could be improved in various aspects, such as preventing social loafing and using cumulative knowledge as feedback. Another possible improvement is to eliminate the patterns that are rated unimportant by the group after the first round. Such pruning will drastically reduce the number of patterns to review from round 2 onward. The smaller set of patterns might allow the experts to evaluate the remaining patterns more carefully and hence improve the final interpretation. Future research should also study the group dynamics for semi-anonymous groups or onymous groups in expert decision-making.

One of the major points of this dissertation is that pattern interpretation is something that should not be done only by KDD researchers or by a single domain expert alone, at least not currently, but a second major point is that computers can greatly facilitate the evaluation process by allowing a group of experts to collaborate on decisions about the discovered patterns. In conclusion, this research has established the benefit of CI for knowledge discovery. The implications of the findings suggest opportunities for further research on CI for knowledge discovery. The CIS constitutes the building blocks for the development of new applications for CI and for expert group collaboration in decision-making.

LIST OF REFERENCES

[1] Ackerman, M. *Answer Garden: A Tool for Growing Organizational Memory.* Alfred P. Sloan School of Management, Massachusetts Institute of Technology, Boston, 1994.

[2] Adomavicius, G. and Tuzhilin, A. *Using Data Mining Methods to Build Customer Profiles.* Computer, 74-82, 2001.

[3] Adriaans, P. and Zantinge, D. *Data Mining*. Addison-Wesley Longman, 1996.

[4] Agrawal, R. and Stolorz, P. (eds.), *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining*. AAAI Press, New York, 1998.

[5] Agrawal R., M.H., Srikant R., Toivonan H., and Verkamo SI. *Fast discovery of association rules.* in Fayyad, U.M., Piatetsky-Shapiro, G., Uthurusamy, R. eds. *Advances in Knowledge Discovery and Data Mining,* MIT Press, Cambridge, MA, 1996, 307-328.

[6] Anastasi, A. *Psychological Testing*. MacMillan Publishing Co., Inc., New York, 1982.

[7] Ayyub, B. *Elicitation of Expert Opinions for Uncertainty and Risks*. CRC Press, New York, 2001.

[8] Baecker, R., Grudin, J., Buxton, W., and Greenberg, S. *From Customizable Systems to Intelligent Agents.* in Baecker, R., Grudin, J., Buxton, W., and Greenberg, S. eds. *Human-Computer Interaction: Toward the year 2000*, Morgan Kaufmann Publishers, Inc., San Francisco, Ca, 1995, 783-792.

[9] Baecker, R., Grudin, J., Buxton, W., and Greenberg, S. Groupware and Computer Supported Cooperative Work. in Baecker, R., Grudin, J., Buxton, W., and Greenberg, S. eds. *Human-Computer Interaction: Toward the year 2000*, Morgan Kaufmann Publishers, Inc, San Francisco, Ca, 1995, 741-753.

[10] Bayardo, R. (ed.), *The Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2000)*. ACM, Boston, 2000.

[11] Bland, J.M. and Altman, D. G. Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. *The LANCET*. 307-310.

[12] Brachman, R. and Anand, T. *The Process of Knowledge Discovery in Databases: A Human-Centered Approach.* in Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. eds. *Advances in Knowledge Discovery and Data Mining*, AAAI/MIT Press, Menlo Park, Ca, 1996, 37-57.

[13] Brossette, S.E. *Data Mining and Epidemiologic Surveillance.* Computer and Information Sciences, University of Alabama at Birmingham, 1998.

[14] Brossette, S.E., Sprague, A.P., Hardin, J.M., Waites, K.B., Jones, W.T., and Moser, S.A. *Association Rules and Data Mining in Hospital Infection Control and Public Health Surveillance.* JAMIA, 5 (July/August), 373-381, 1998.

[15] Cockburn, A. and Thimbleby, H. *A Reflexive Perspective of CSCW.* SIGCHI Bulletin, 23 (3). 63-68, 1991.

[16] Dean, A.G., Fagan, R. F., and Panter-Connah, B. J. *Computerizing Public Health Surveillance Systems.* in Teutsch, S.M. and Churchill, R.E. eds. *Principles and Practice of Public Heallth Surveillance.* Oxford University Press, New York, 1994, 200-217.

[17] Delbecq, A., Van De Ven, A., and Gustafson, D. H. *Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes.* Scott, Foresman, Chicago, IL, 1975.

[18] Emori, T.G., Culver, D. H., Horan, T. C., Jarvis, W. R., White, J. W., Olson, D. R., Banerjee, S., Edwards, J., Martone, W. J., Gaynes, R. P., Hughes, J. M. *National nosocomial infections surveillance system (NNIS): Description of surveillance methods.* American Journal of INFECTION CONTROL, 19 (1). 19-35, 1991.

[19] Farrington, C.P., Andrews, N.J., Beale, A.D., and Catchpole, M.A. *A Statistical Algorithm for the Early Detection of Outbreaks.* Journal of the Royal Statistical Society *A* (159. Part 3). pp. 547-563, 1996.

[20] Fayyad, U., Chaudhuri, S., and Madigan, D. (eds.), *The ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-99).* ACM Press, San Diego, 1999.

[21] Fayyad, U.M. and Uthurusamy, R. (eds.), *The First International Conference on Knowledge Discovery & Data Mining.* AAAI Press, 1995.

[22] Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. *From Data Mining to Knowledge Discovery: An Overview.* AI Magazine, 17 (3). 37-54, 1996.

[23] Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. (eds.), *Advances in Knowledge Discovery and Data Mining.* AAAI/MIT Press, Cambridge, MA, 1996.

[24] Groth-Marnat, G. *Handbook of Psychological Assessment*. John Wiley & Sons, Inc., New York, 18-20, 1999.

[25] Grudin, J. *Groupware and Social Dynamics: Eight Challenges for Developers.* Comunications of the ACM, *37* (1), 97-105, 1994.

[26] Grudin, J., *Why CSCW Applications Fail: Problems in Design and Evaluation.* CSCW, ACM Press, 85-93, 1988.

[27] Hand, D., Keim, D., and Ng, R. (eds.), *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2002.

[28] Heckerman, D., Mannila, H., and Uthurusamy, R. (eds.), *The Third International Conference on Knowledge Discovery & Data Mining*. AAAI Press, 1997.

[29] Hiltz, S. R., Dufner, D., Fjermestad, J., Kim, Y., Ocker, R., Rana, A., and Turoff, M. *Distributed Group Support Systems: Theory Development and Experimentation*. in Olsen, B.M., Smith, J.B., and Malone, T. eds. *Coordination Theory and Collaboration Technology*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1996.

[30] Jones, R.N. *The current and future impact of antimicrobila resistance among nosocomial bacterial pathogens*. Diagn. Microbiol. Infect. Dis., 15. 3-10, 1992.

[31] Julisch, K. and Dacier, M., *Mining Intrusion Detection Alarms for Actionable Knowledge*. in KDD - 2002, ACM, 366 - 375, 2002.

[32] Koontz, F.P. *A review of traditional resistance surveillance methodologies and infection control*. Diagn. Microbiol. Infect. Dis., *15*. 43-47.

[33] Latane, B., Williams, K., and Harkins, S. *Many hands make light the work: The causes and consequences of social loafing*. Journal of Personality and Social Psychology, *37*. 822-832, 1979.

[34] Livingston, G.R., Rosenberg, J. M., and Buchanan, B. G, *Closing the Loop: Heuristics for Autonomous Discovery*. 2001 IEEE International Conference on Data Mining, IEEE Computer Society Press, 393-400, 2001.

[35]. Martino, J. *Technological Forecasting For Decision Making*. McGraw-Hill Engineering and Technology Management Series, New York, 1993.

[36] Matheus, C.J., Piatetsky-Shapiro, G., and McNeill, D. *Selecting and Reporting What Is Interesting*. in Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. eds. *Advances in Knowledge Discovery and Data Mining*, AAAI/MIT Press, Menlo Park, Ca, 495-515, 1996.

[37] Mayhew, D.J. *Principles and Guidelines in Software Interface Design*. Prentice Hall, Englewood Cliffs, New Jersey, 1992.

[38] Meyer, G., Finn, S., Eyde, L., Kay, G., Moreland, K., Dies, R.,Eisman, E., Kugiszyn, and T., Reed, G. *Psychological Testing and Psychological Assessment, A Review of Evidence and Issues*. American Psychologist, 56 (2). 128-165, 2001.

[39] Huynh, M., Popkin, L., and Stecker, M. *Constructing a Corporate Memory Infrastructure from Internet Discovery Technologies*. http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/CorInfSys/huynh/cmi.html.

[40] MMWR. *Guidelines for Evaluating Surveillance Systems*, Centers For Disease Control, 1988.

[41] MMWR. *Public health focus surveillance: prevention and control of nosocomial infections.*, Centers For Disease Control, 783-787, 1992.

[42] Neisser, U., Boodoo, G., Bouchard, T. Jr., Boykin, A. W., Brody, N., Ceci, S., Halpern, D., Loehlin, J., Perloff, R., Sternberg, R., and Urbina, S. *Intelligence: Knowns and Unknowns*. American Psychologist, 51 (2). 77-102, 1996.

[43] Neu, H.C., R.J. Duma, R.N. Jones, J.E., McGowan, J., O'Brien, T.F., Sabath, L. D. and et.al. *Antibiotic resistance: epidemiology and therapeutics.* Diagn.Microbiol.Infect.Dis., *2*. 53-60, 1992.

[44] Ng, R. (ed.), *1996 SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery*. Dept. of Computer Science, The University of British Columbia, Vancouver, B.C., 1996.

[45] Nielsen, J. *Usability Engineering*. AP Professional, Boston, 1993.

[46] Orlikowski, W., *Learning From Notes: Organizational Issues in Groupware Implementation*. CSCW, ACM Press, 1992.

[47] Piatetsky-Shapiro, G. and Frawley W. *Knowledge Discovery in Databases.* AAAI Press / MIT Press, Cambridge, MA, 1991.

[48] Price, W.T. (ed.), *Data Mining Technology and Applications in the Government Community*. AFCEA International, Fairfax, 1999.

[49] Provost, F., and Srikant, R. (eds.), *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2001.

[50] Reilly, K.D. *The Delphi Technique Fundamentals and applications*. in Borko, H. ed. *Targets for Research in Library Education*, American Library Assoc., Chicago, IL, 187-189, 1973.

[51] Saaty, T.L. *The Analytic Hierarchy Process*. McGraw Hill Company, New York, 1980.

[52] Salton, G., McGill *Introduction to Modern Information Retrieval*. McGraw-Hill, 1983.

[53] Scheibe, M., Skutsch, M., Schofer, J. *Experiments in Delphi Methodology.* Linstone, H. and Turoff, M. eds. *The Delphi Method: Techniques and Applications*, Addison-Wesley, Reading, MA, 262-287, 1975.

[54] Shlaes, D.M., Gerding, D.N., John, J.F., Craig, W.A., Bornstein, D.L., and Duncan, R.A. *Society of American Joint Committeee on the Prevention of Antimicrobial Resistance: Guidelines for the prevention of antimicrobial resistance in hospitals.* Clin Infect Dis, 25. 584-599, 1997.

[55] Shneiderman, B. *Designing the User Interface - Strategies for Effective Human Computer Interaction.* Addison Wesley, 1998.

[56] Siegel, S., and Castellan, N.J. JR *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, 1988.

[57] Silberschatz, A., and Tuzhilin, A. *User-Assisted Knowledge Discovery: How Much Should the User Be Involved.* 1996 SIGMOD Workshop on Research Issues on Data Mining and Knowledge Discovery. The University of British Columbia, 89-94, 1996.

[58] Simoudis, E., Han, J., and Fayyad, U. M. (eds.), *The Second International Conference on Knowledge Discovery & Data Mining*. AAAI Press, 1996.

[59] Simoudis, E., Livezey, B., and Kerber, R. *From Data Mining to Knowledge Discovery: Current Challenges and Future Directions*. in Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. eds. *Advances in Knowledge Discovery and Data Mining*, AAAI/MIT Press, Menlo Park, Ca, 353-373, 1996.

[60] Singh, I., Stearns, B., and Johnson, M., and Enterprise Team, *Designing Enterprise Applications with the $J2EE_{TM}$ Platform*. Sun Microsystems, Inc., 2002.

[61] Smith, J. *Collective intelligence in Computer-Based Collaboration*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1994.

[62] Stedman, C. *Data Mining for Fool's Gold.* Computerworld, 1997.

[63] Turoff, M. *Computer-Mediated Communication Requirements for Group Support.* Journal of Organizational Computing, 1. 85-113, 1991.

[64] Turoff, M., and Hiltz, S.R. *Computer Based Delphi Process*. in Adler, M. and Ziglio, E. eds. *Gazing Into the Oracle: The Delphi Method and Its Application to Social Policy and Public Health*, Jessica Kingsley Publisher, London, 1995.

[65] Uebersax, J. *Statistical Methods for Rater Agreement.* http://ourworld.compuserve.com/homepages/jsuebersax/agree.htm, 2001.

[66] Uthurusamy, R. *From Data Mining to Knowledge Discovery: Current Challenges and Future Directions.* in Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. eds. *Advances in Knowledge Discovery and Data Mining*, AAAI/MIT Press, Menlo Park, Ca, 561-569, 1996.

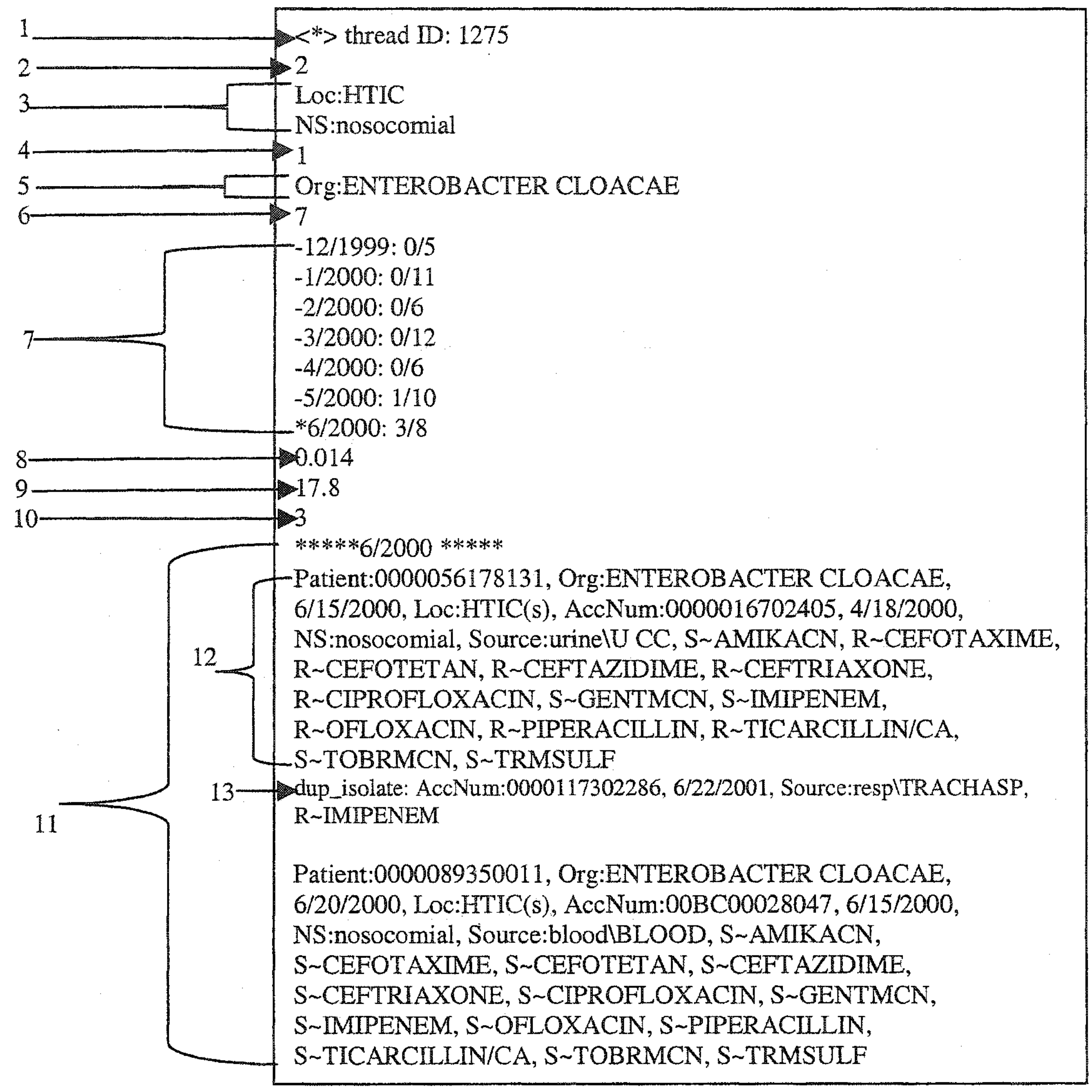
[67] Weiss, S.M., Verma, Naval K., *A System for Real-time Competitive Market Intelligence.* in KDD - 2002, ACM, 360-365, 2002.

[68] Wong, D.Y., Jones, W.T., Brossette, S.E., Hardin, J.M., and Moser, S.A. *A Strategy for Geomedical Surveillance Using the Hawkeye Knowledge Discovery System.* in Geomed '97, Teubner-Verlag, 204-213, 1997.

[69] Wright, P., Butler, C., and Ray, J. *Using Knowledge Discovery to Identify Similar Bridges.* in Federal Data Mining Symposium & Exposition '99, AFCEA International, 53-60, 1999.

APPENDIX A

AN EXAMPLE DMSS OUTPUT PATTERN



Explanation of the data fields in the example pattern output from DMSS:

1. **Thread ID**

    Identifies a pattern within a month's data. However, when the same pattern appears again in the past or in future months, the thread ID may be different from the one assigned for this month. Therefore, thread ID is not a unique identifier for a pattern that occurs in more than one month.

2. **Surveillance attribute count**

Indicates the number of surveillance attributes to follow.

3. **Surveillance attribute list**

List of attributes in the surveillance group of this pattern, one attribute is listed on each line. Each attribute is composed of a prefix and a suffix, separated by a ':'. The prefix is the attribute category; the suffix is the attribute label (i.e. name). e.g. Loc:HTIC means the attribute is a location, and the location name is HTIC. A list of attribute categories is attached at the end of this document.

4. **Outcome attribute count**

Indicates the number of outcome attributes to follow.

5. **Outcome attribute list**

List of attributes in the outcome group of this pattern, one attribute is listed on each line. Explanation from *Surveillance attributes* above applies here.

6. **Surveillance period's month count**

Indicates the number of months in this pattern's surveillance period from this month's data mining.

7. **Surveillance period's month list**

List each month and the associated isolate counts in this pattern's surveillance period, one month per line. Format: -5/2000: 1/10 or *6/2000: 3/8 The first part is the date (MM/YYYY) or (M/YYYY). Colon is the separator. The ratio that follows is interpreted as follow: the numerator is the no. of isolates in the outcome group, the denominator is the no. of isolates in the surveillance.

Basically, 1/10 means one out of the ten isolates tested in that month has the characteristics of the pattern.

For the pattern in the example, *6/2000: 3/8 means in June 2000, 3 out of the 8 nosocomial isolates from HTIC is tested positive for the organism ENTEROBACTER CLOACAE.

The '*' indicates that the ratio of that month is the "front window" in the data-mining algorithm. There may be multiple months with a '*' as prefix. Pattern 2 below is such example. When there are multiple '*' ratios, the last one on the list (i.e. the month the data mining is performed for) is the "focus month" and the others are the supporting months. The no. of front window months in a pattern may range from 1 to 3.

The '-' indicates that the ratio of that month is in the "back window"in the algorithm. The number of back window months may range from 6 to 9

8. p-value

This is one of the statistics from the data-mining algorithm.

9. Relative difference

This is one of the statistics from the data-mining algorithm.

10. Patient isolate count

The number of patients in the total outcome group isolates in the front windows ratios of the pattern. A bacterial isolate is a test sample taken from a specimen (e.g. urine, blood) obtained from a patient.

11. Patient isolate list

List the details of each isolates in the outcome group of the front window ratios and the test results. Some months may not have any isolates. The isolates are grouped by month in the listing. For a pattern that appears in other months in a relatively close time period (e.g. within 3 months), the same isolate might be listed again under that same month's ratio.

12. A bacterial isolate from a patient

An isolate listing contains the following:

a. Patient:0000056178131 – patient identification no. It's unique for each patient regardless how many times the patient has been admitted.
b. Org:ENTEROBACTER CLOACAE – organism found in the isolate
c. The date following organism (e.g. 6/15/2000) – specimen date
d. Loc:xxx (s) – location where the patient was when the specimen was taken. (s) indicates patient was still in the hospital. The other option is (d) for discharged.
e. AccNum:0000016702405 – accession number of the primary isolate
f. 4/18/2000 – date the patient was admitted to the location
g. NS:nosocomial – indicates the origin of the infection. The choices are nosocomial, community acquired, and unknown.
h. Source:urine\U CC - source of the specimen from the patient
i. Following source is the antibiogram (test result) of the isolate. The antibiogram indicates whether the organism identified is susceptible (S~), or resistant (R~), or intermediate (I~) to the antibiotics. e.g. S~AMIKACN means the bacteria is susceptible to the drug AMIKACN. The number of drug test results in the antibiogram depends on what the data-mining algorithm digs up. An example antibiogram is: S~AMIKACN,

R~CEFOTAXIME, R~CEFOTETAN, R~CEFTAZIDIME,

R~CEFTRIAXONE, R~CIPROFLOXACIN, S~GENTMCN,

S~IMIPENEM, R~OFLOXACIN, R~PIPERACILLIN,

R~TICARCILLIN/CA, S~TOBRMCN, S~TRMSULF

13. Duplicate isolates

An isolate may have zero or more duplicate isolates. Each duplicated isolate is listed on one line. A duplicate isolate listing contains the following:

a. AccNum:0000117302286 – accession number of the duplicate isolate
b. 6/22/2001 – specimen date of the duplicate isolate
c. Source:resp\TRACHASP - source of the specimen from the patient
d. Antibiogram, e.g.R~IMIPENEM. The antibiogram of the duplicate isolate only contains the test results that are different from the primary isolate.

Important notes:

1. A pattern is identified by the attributes in its outcome and surveillance attribute list. A pattern may appear again a few months later.
2. All occurrences of a pattern form the history of the pattern in CIS.

APPENDIX B

EXPERIMENT DATA – RATINGS OF EACH MONTH

April 2001 Ratings

	E1			E2			E3			E4			E5			E6			Group			C
ID	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	
442	3	3	3	4	4	4	2	2	2	3	3	3	4	4	4	1	3	3	2.8	3.1	3.1	3
475	2	2	2	1	1	1	4	4	4	4	4	4	4	4	4	4	4	4	3.1	3.1	3.1	4
660	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.1	1.1	1.1	1
732	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.3	1.3	1.3	1
772	3	3	3	3	3	3	4	4	4	3	3	3	2	2	2	1	3	3	2.6	3	3	4
792	3	3	3	4	3	3	2	2	2	2	2	2	2	2	2	3	3	3	2.6	2.5	2.5	4
1149	1	1	1	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	1.3	1.3	1.3	1
1155	2	2	2	2	2	2	3	3	3	1	1	1	3	3	2	2	2	2	2.1	2.1	2	1
1233	2	2	2	3	3	3	2	2	2	2	2	2	1	1	1	3	3	3	2.1	2.1	2.1	2
1246	4	4	4	3	3	3	2	2	2	1	1	1	5	5	5	4	4	4	3.1	3.1	3.1	4
1263	2	2	2	4	4	4	2	2	2	1	1	1	1	2	2	4	4	4	2.3	2.5	2.5	2
1312	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	3	2	2	2	1.8	1.8	2
1708	2	2	2	3	3	3	3	3	3	2	2	2	2	2	2	2	2	2	2.3	2.3	2.3	2
2039	2	2	2	1	1	1	3	3	3	3	3	3	2	2	2	2	2	2	2.1	2.1	2.1	4
2086	2	2	2	4	4	4	4	4	4	3	4	4	1	1	1	2	2	2	2.6	2.8	2.8	3
2096	2	2	2	4	3	3	3	3	3	3	3	3	1	1	1	1	1	1	2.3	2.1	2.1	1
2111	2	2	2	1	3	3	3	3	3	3	3	3	2	2	2	3	3	3	2.3	2.6	2.6	2
2133	2	2	2	1	1	1	3	3	3	2	2	2	2	2	2	2	2	2	2	2	2	1
2139	3	3	3	2	2	2	3	3	3	1	1	1	2	2	2	1	1	1	2	2	2	5
2164	2	2	2	4	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2.6	2.5	2.5	2
2201	4	4	4	3	3	3	4	4	4	4	4	4	1	1	2	3	3	3	3.1	3.1	3.3	4
2203	2	2	2	4	4	4	3	3	3	4	4	4	1	1	1	4	4	4	3	3	3	3
2230	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2234	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	4	1	1	1.5	1	1	1
2236	4	4	4	2	2	2	2	2	2	1	1	1	1	1	1	2	2	2	2	2	2	1
2245	2	2	2	1	1	1	1	1	1	1	1	1	2	2	2	1	1	1	1.3	1.3	1.3	1
2266	4	4	4	3	3	3	2	2	2	1	1	1	1	1	1	1	1	1	2	2	2	1
2278	2	2	2	4	4	4	4	4	4	3	3	3	1	1	1	3	3	3	2.8	2.8	2.8	3
2286	2	2	2	1	1	1	2	2	2	4	4	4	1	1	1	1	3	3	1.8	2.1	2.1	2
2312	5	5	5	3	4	4	4	4	4	4	4	4	5	5	5	4	4	4	4.1	4.3	4.3	4
2316	2	2	2	3	3	3	4	4	4	3	3	3	2	2	2	1	3	3	2.5	2.8	2.8	4
defaults:	4	4	4	8	7	7	6	6	6	9	9	9	16	15	14	12	8	8				11
Changes:		0	0		5	0		0	0		1	0		1	2		6	0		3.3	0.3	

Number of patterns = 31. E1 – E6 are the in-group experts. C is the criterion rating. r1-r3 are the rounds.

May 2001 Ratings

	E1			E2			E3			E4			E5			E6			E7			Group			C
ID	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	
475	2	2	2	4	3	3	3	3	3	5	5	5	1	2	2	2	2	2	3	3	3	2.8	2.8	2.8	4
741	4	2	2	1	1	1	2	2	2	1	1	1	1	1	1	2	2	2	1	1	1	1.8	1.5	1.5	3
1155	3	3	3	1	1	1	4	4	4	1	3	3	3	3	3	2	2	2	4	4	4	2.3	2.6	2.6	3
1312	2	2	2	1	1	1	3	3	3	3	3	3	1	1	1	4	4	4	1	1	1	2.3	2.3	2.3	2
2230	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2316	4	4	4	4	4	4	4	4	4	2	4	4	2	2	3	4	4	4	5	5	5	3.3	3.6	3.8	3
6226	4	4	4	1	1	1	4	4	4	3	4	4	1	1	1	2	4	4	4	4	4	2.5	3	3	2
6227	1	3	3	3	3	3	3	3	3	1	3	3	3	3	3	4	4	4	3	3	3	2.5	3.1	3.1	4
6232	3	3	3	1	1	1	3	3	3	2	2	2	1	1	1	2	2	2	4	4	4	2	2	2	2
6234	2	2	2	1	1	1	3	3	3	2	2	2	1	1	1	2	2	2	4	4	4	1.8	1.8	1.8	2
6245	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6246	2	2	2	3	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.6	1.5	1.5	2
6247	2	2	2	1	1	1	2	2	2	3	3	3	1	1	1	1	1	1	3	3	3	1.6	1.6	1.6	1
6248	3	3	3	3	2	2	2	2	2	1	1	1	1	1	1	1	1	1	3	3	3	1.8	1.6	1.6	1
6249	4	4	4	1	2	2	3	3	3	1	3	3	3	3	3	4	4	4	4	4	4	2.6	3.1	3.1	3
6250	4	4	4	3	3	3	3	3	3	4	4	4	1	2	2	3	3	3	2	2	2	3	3.1	3.1	2
6251	2	2	2	1	1	1	3	3	3	4	4	4	1	1	1	3	3	3	2	2	2	2.3	2.3	2.3	5
6252	2	2	2	2	2	2	3	3	3	1	1	1	1	1	1	4	3	3	3	3	3	2.1	2	2	5
6253	2	2	2	3	3	3	3	3	3	3	3	3	1	1	1	1	1	1	3	3	3	2.1	2.1	2.1	1
6254	2	2	2	1	1	1	3	3	3	1	1	1	5	5	5	4	4	4	5	3	3	2.6	2.6	2.6	4
6255	2	2	2	2	2	2	3	3	3	1	1	1	5	5	5	4	5	5	5	3	3	2.8	3	3	4
6256	5	5	5	4	4	4	1	3	3	1	1	1	5	5	5	4	5	5	5	5	5	3.3	3.8	3.8	4
defaults	3	2	2	12	11	11	3	2	2	12	9	9	15	13	13	6	6	6	5	5	5				5
Changes:		2	0		4	0		1	0		5	0		2	1		4	0	0	2	0		3.7	0.2	

Number of patterns = 22. Number of patterns = 31. E1 – E6 are the in-group experts. C is the criterion rating. r1-r3 are the rounds. E7 is the out-group expert.

June 2001 Ratings

	E1			E2			E3			E4			E5			E6			E7			Group			C
ID	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	
741	2	2	2	3	3	3	2	2	2	3	3	3	1	1	1	1	1	1	2	2	2	2	2	2	3
1312	2	2	2	2	2	2	3	3	3	4	4	4	1	1	1	4	4	4	2	2	2	2.6	2.6	2.6	2
2316	1	1	1	1	1	1	3	3	3	1	1	1	3	2	2	2	1	1	1	2	2	1.8	1.5	1.5	3
6051	4	4	4	1	5	5	1	5	5	1	1	5	5	5	5	5	5	5	4	4	4	2.8	4.1	4.8	4
6246	1	1	1	1	1	1	3	3	3	2	2	2	1	1	1	1	1	1	1	1	1	1.5	1.5	1.5	1
6251	1	1	2	1	1	1	3	3	3	3	3	3	1	1	1	4	3	3	3	5	4	2.1	2	2.1	2
6255	3	3	3	1	2	2	2	2	2	1	1	1	3	3	3	3	3	3	3	3	3	2.1	2.3	2.3	2
6266	2	2	2	1	2	2	2	2	2	1	1	1	3	3	3	2	2	2	1	1	1	1.8	2	2	1
6267	1	1	1	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1.1	1.1	1.1	2
6268	2	2	2	1	2	2	3	3	3	2	2	2	4	4	4	1	1	1	1	1	1	2.1	2.3	2.3	3
6269	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1	3	1	1	1	1	1	1.5	1.1	1.1	1
6270	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.3	1.3	1.3	1
6271	2	2	2	1	1	1	2	2	2	2	2	2	3	2	2	2	2	2	1	1	1	2	1.8	1.8	1
6272	4	4	4	1	1	1	3	3	3	1	1	1	4	4	4	3	3	4	2	2	2	2.6	2.6	2.8	2
6273	2	2	2	1	2	2	4	4	4	4	4	4	3	3	3	2	2	2	1	1	1	2.6	2.8	2.8	2
6274	3	3	3	1	1	1	1	1	1	2	2	2	3	3	3	2	2	2	1	1	1	2	2	2	2
6275	3	3	3	1	2	2	4	4	4	2	2	2	4	3	3	2	2	2	2	3	3	2.6	2.6	2.6	1
6276	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	2	1	1	1	1	1	1.5	1.3	1.3	1
6277	1	1	1	1	1	1	3	3	3	1	1	1	2	2	2	3	3	3	3	3	3	1.8	1.8	1.8	1
6278	3	3	3	1	2	2	5	5	5	1	1	2	2	2	2	4	4	5	3	3	3	2.6	2.8	3.1	5
6279	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6280	3	3	3	1	2	2	1	1	1	1	1	1	2	2	2	3	2	2	4	4	4	1.8	1.8	1.8	3
6281	3	3	3	1	4	4	4	4	4	4	4	4	4	4	4	4	4	4	3	3	3	3.3	3.8	3.8	2
6282	1	1	1	1	1	1	2	2	2	3	3	3	1	1	1	1	1	1	4	4	4	1.5	1.5	1.5	1
6283	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	4	4	4	1.1	1.1	1.1	1
6284	1	1	1	1	1	1	3	3	3	1	1	1	1	1	1	2	1	1	1	1	1	1.5	1.3	1.3	1
6285	5	5	5	1	5	5	4	4	4	2	5	5	3	3	4	5	5	5	5	5	5	3.3	4.5	4.6	3
defaults	10	10	9	25	15	15	5	4	4	14	14	12	12	12	12	8	8	8	13	12	12				12
Changes:		0	1		10	0		1	0		1	2		3	1		6	2	0	3	1		5.4	1.4	

Number of patterns = 27. Number of patterns = 31. E1 – E6 are the in-group experts. C is the criterion rating. r1-r3 are the rounds. E7 is the out-group expert.

July 2001 Ratings

	E1			E2			E3			E4			E5			E6			E7			Group			C
ID	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	
6039	4	4	4	1	4	4	3	3	3	3	3	3	3	3	3	4	4	4	5	5	5	3	3.5	3.5	5
6051	1	1	3	1	3	3	4	4	4	1	1	1	3	3	3	4	4	4	2	5	4	2.3	2.6	3	4
6215	4	4	4	1	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	3.5	4	4	4
6255	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1	3	3	2	2	2	2	1.5	1.5	1.3	2
6267	1	1	1	1	1	1	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1.3	1.3	1.3	2
6268	4	4	4	1	3	3	2	2	2	4	4	4	4	4	4	4	4	4	4	4	4	3.1	3.5	3.5	4
6270	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.3	1.3	1.3	1
6271	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1	2	2	1	1	1	1	1.3	1.3	1.1	1
6285	2	2	2	1	2	2	4	4	4	1	1	1	3	3	3	3	3	3	4	5	4	2.3	2.5	2.5	2
6291	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.3	1.3	1.3	1
6292	2	2	2	1	1	1	3	3	3	1	1	1	1	1	1	1	1	1	1	1	1	1.5	1.5	1.5	2
6293	3	3	3	1	2	2	1	3	3	3	3	3	3	2	3	3	3	3	3	3	3	2.3	2.6	2.8	2
6294	2	2	2	1	2	2	3	3	3	2	2	2	1	1	1	3	3	2	1	1	1	2	2.1	2	2
6295	2	2	2	1	2	2	3	3	3	1	1	1	1	1	1	1	1	1	1	1	1	1.5	1.6	1.6	2
6296	1	1	1	2	2	2	1	1	1	1	1	1	2	2	2	2	2	2	1	1	1	1.5	1.5	1.5	3
6297	1	1	1	1	2	2	3	3	3	3	3	3	3	3	3	2	3	3	1	1	1	2.1	2.5	2.5	2
6298	1	1	1	1	3	2	3	3	3	4	4	4	1	1	1	1	1	1	1	1	1	1.8	2.1	2	2
6299	1	3	3	1	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4	2.8	3.8	3.8	4
6300	2	2	2	1	2	2	2	2	2	3	3	3	2	2	2	1	1	2	3	3	3	1.8	2	2.1	2
6301	1	1	1	1	1	1	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1.3	1.3	1.3	1
6302	1	1	1	1	1	1	3	3	3	3	3	3	1	1	1	1	1	1	3	3	3	1.6	1.6	1.6	2
6303	4	4	4	1	5	5	4	4	4	4	4	4	4	4	4	3	4	4	3	5	5	3.3	4.1	4.1	4
6304	2	2	2	1	1	2	2	2	2	2	2	2	2	2	1	1	1	2	4	3	3	1.6	1.6	1.8	1
6305	1	1	1	1	1	1	3	3	3	2	2	2	1	1	1	1	1	1	4	4	4	1.5	1.5	1.5	2
6306	1	1	1	1	1	1	2	2	2	3	3	3	1	1	1	1	1	1	4	3	3	1.5	1.5	1.5	2
6307	1	1	1	4	4	4	4	4	4	2	4	4	1	2	2	1	1	1	4	4	4	2.1	2.6	2.6	2
defaults	13	12	11	24	11	10	2	1	1	9	9	9	14	13	13	13	13	11	11	11	11				5
Changes:		1	1		13	2		1	0		1	0		2	2		3	5	0	5	2		5.6	1.5	

Number of patterns = 26. Number of patterns = 31. E1 – E6 are the in-group experts. C is the criterion rating. r1-r3 are the rounds. E7 is the out-group expert.

August 2001 Ratings

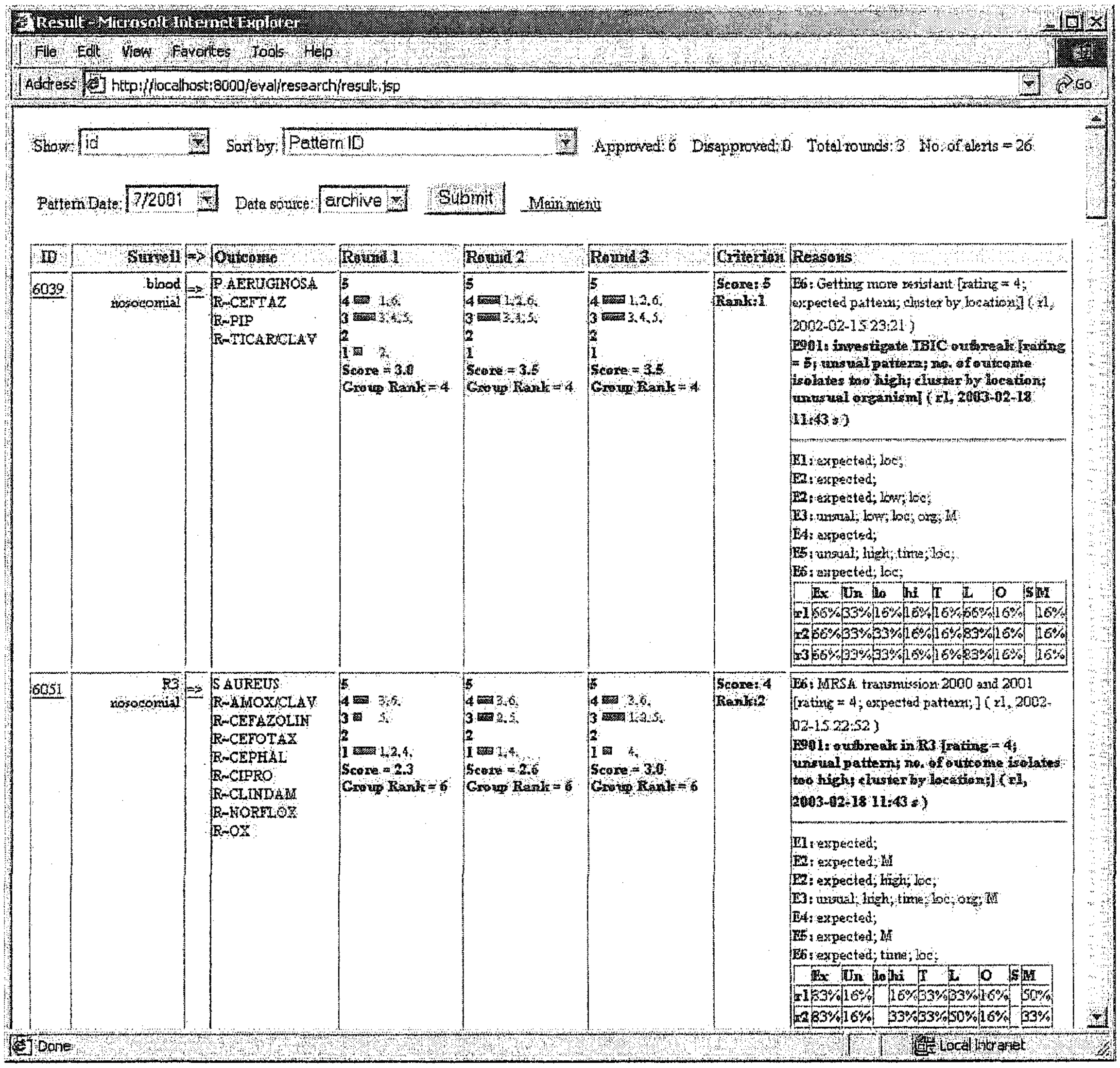
	E1			E2			E3			E4			E5			E6			E7			Group			C
ID	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	r1	r2	r3	
6040	3	3	3	1	2	2	3	3	3	2	2	2	3	2	2	1	2	2	5	5	5	2.1	2.3	2.3	1
6051	3	3	3	1	3	3	3	3	3	1	3	3	3	3	3	5	5	5	4	4	4	2.6	3.3	3.3	4
6269	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.1	1.1	1.1	1
6291	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.3	1.3	1.3	1
6292	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.3	1.3	1.3	2
6298	2	2	2	1	1	1	2	2	2	3	1	1	1	1	1	1	1	1	1	1	1	1.6	1.3	1.3	2
6302	2	2	2	1	1	1	2	2	2	2	2	2	1	1	1	2	2	2	1	1	1	1.6	1.6	1.6	2
6303	3	3	3	1	3	3	4	4	4	4	4	4	3	1	1	1	1	1	1	3	3	2.6	2.6	2.6	2
6323	4	4	4	1	4	4	4	4	4	1	3	3	4	4	4	3	4	4	4	4	4	2.8	3.8	3.8	5
6324	2	2	2	1	3	3	4	4	4	1	2	2	3	2	1	5	3	3	1	1	1	2.6	2.6	2.5	3
6325	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.3	1.3	1.3	1
6326	2	2	2	1	1	1	3	3	3	2	2	2	1	1	1	2	2	2	3	3	3	1.8	1.8	1.8	3
6327	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1.1	1.1	1	1
6328	2	2	2	1	1	1	2	2	2	2	2	2	1	1	1	2	2	2	1	1	1	1.6	1.6	1.6	1
6329	3	3	3	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.5	1.5	1.5	1
6330	3	3	2	1	2	2	4	4	4	1	1	1	1	1	1	1	1	1	2	2	2	1.8	2	1.8	2
6331	3	3	3	1	3	3	4	4	4	3	3	3	4	4	4	2	3	3	4	4	4	2.8	3.3	3.3	4
6332	3	3	3	1	4	4	4	4	4	1	3	3	4	4	4	4	4	4	1	1	1	2.8	3.6	3.6	4
6333	3	3	3	1	4	4	4	4	4	3	3	3	4	4	4	3	5	5	3	4	4	3	3.8	3.8	4
6334	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.1	1.1	1.1	1
6335	3	3	3	1	2	3	4	4	4	1	1	1	2	1	1	3	3	3	3	3	3	2.3	2.3	2.5	2
6336	2	2	2	1	2	2	3	3	3	2	2	2	1	1	1	3	2	2	5	5	3	2	2	2	1
6337	2	2	2	1	1	1	3	3	3	2	2	2	1	1	1	2	2	2	2	2	2	1.8	1.8	1.8	2
6338	3	3	3	1	2	2	4	4	4	3	3	3	2	2	2	3	3	3	3	3	3	2.6	2.8	2.8	3
6339	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	2	2	2	3	3	3	1.5	1.5	1.5	2
6340	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1	2	2	2	1	1	1	1.3	1.3	1.3	2
6341	1	1	1	1	1	1	3	3	3	3	3	3	1	1	1	2	2	2	4	4	4	1.8	1.8	1.8	2
6342	2	2	2	1	1	1	3	3	3	1	1	1	1	1	1	2	2	2	2	2	2	1.6	1.6	1.6	1
6343	2	2	2	1	2	2	4	4	4	4	4	4	1	1	1	2	2	2	2	2	2	2.3	2.5	2.5	1
6344	2	2	2	1	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1.3	1.3	1.3	1
6345	2	2	2	1	1	1	2	2	2	2	2	2	1	1	1	2	2	2	1	1	1	1.6	1.6	1.6	1
6346	3	3	3	1	3	3	3	3	3	4	4	4	1	1	2	4	4	4	2	2	3	2.6	3	3.1	2
6347	3	3	3	1	3	3	3	3	3	4	4	4	1	2	2	3	3	3	2	2	3	2.5	3	3	1
6348	2	2	2	1	2	2	2	2	2	4	4	4	1	1	1	3	3	2	2	2	2	2.1	2.3	2.1	1
6349	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	4	3	3	2	2	2	1.6	1.5	1.5	1
6350	3	3	3	1	4	4	3	3	3	4	4	4	1	1	2	3	3	3	2	2	2	2.5	3	3.1	1
6351	3	3	3	1	3	3	1	1	3	4	3	3	1	2	2	3	3	3	4	4	4	2.1	2.5	2.8	1
6352	2	2	2	1	1	1	3	3	3	3	3	3	1	1	1	2	2	2	1	1	1	2	2	2	2
6353	2	2	2	1	2	2	3	3	3	3	3	3	1	1	1	3	3	3	1	1	1	2.1	2.3	2.3	2

6354	1	1	1	1	1	1	2	2	2	4	4	4	1	1	1	2	2	2	5	5	5	1.8	1.8	1.8	1
defaults	5	5	5	40	21	21	3	3	2	18	14	14	30	28	26	12	11	11	18	17	17				19
Changes:		0	2		19	1		0	1		6	0		6	3		7	1	0	2	3		7.2	1.3	

Number of patterns = 40. Number of patterns = 31. E1 – E6 are the in-group experts. C is the criterion rating. r1-r3 are the rounds. E7 is the out-group expert.

APPENDIX C

CIS RESEARCHER USER INTERFACE SNAP SHOT



Result - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://localhost:8000/eval/research/result.jsp

Show: id Sort by: Pattern ID Approved: 6 Disapproved: 0 Total rounds: 3 No. of alerts = 26

Pattern Date: 7/2001 Data source: archive Submit Main menu

ID	Surveil	=>	Outcome	Round 1	Round 2	Round 3	Criterion	Reasons
6039	blood nosocomial	=>	P AERUGINOSA R-CEFTAZ R-PIP R-TICAR/CLAV	5 4 1,6, 3 3,4,5, 2 1 2, **Score = 3.0** **Group Rank = 4**	5 4 1,2,6, 3 3,4,5, 2 1 **Score = 3.5** **Group Rank = 4**	5 4 1,2,6, 3 3,4,5, 2 1 **Score = 3.5** **Group Rank = 4**	**Score: 5** **Rank:1**	**E6**: Getting more resistant [rating = 4; expected pattern; cluster by location;] (r1, 2002-02-15 23:21) **E901: investigate TBIC outbreak [rating = 5; unsual pattern; no. of outcome isolates too high; cluster by location; unusual organism] (r1, 2003-02-18 11:43 s)** **E1**: expected; loc; **E2**: expected; **E2**: expected; low; loc; **E3**: unsual; low; loc; org; M **E4**: expected; **E5**: unsual; high; time; loc; **E6**: expected; loc;
6051	R3 nosocomial	=>	S AUREUS R-AMOX/CLAV R-CEFAZOLIN R-CEFOTAX R-CEPHAL R-CIPRO R-CLINDAM R-NORFLOX R-OX	5 4 3,6, 3 5, 2 1 1,2,4, **Score = 2.3** **Group Rank = 6**	5 4 3,6, 3 2,5, 2 1 1,4, **Score = 2.6** **Group Rank = 6**	5 4 2,6, 3 1,3,5, 2 1 4, **Score = 3.0** **Group Rank = 6**	**Score: 4** **Rank:2**	**E6**: MRSA transmission 2000 and 2001 [rating = 4; expected pattern;] (r1, 2002-02-15 22:52) **E901: outbreak in R3 [rating = 4; unsual pattern; no. of outcome isolates too high; cluster by location;] (r1, 2003-02-18 11:43 s)** **E1**: expected; **E2**: expected; M **E2**: expected; high; loc; **E3**: unsual; high; time; loc; org; M **E4**: expected; **E5**: expected; M **E6**: expected; time; loc;

6039:

	Ex	Un	lo	hi	T	L	O	S	M
r1	66%	33%	16%	16%	16%	66%	16%		16%
r2	66%	33%	33%	16%	16%	83%	16%		16%
r3	66%	33%	33%	16%	16%	83%	16%		16%

6051:

	Ex	Un	lo	hi	T	L	O	S	M
r1	83%	16%		16%	33%	33%	16%		50%
r2	83%	16%		33%	33%	50%	16%		33%

Done Local intranet

APPENDIX D

MID-STUDY AND POST-STUDY SURVEY RESPONSE SUMMARY

Mid-study survey response summary

	Features	Not at all 1	2	3	4	Always 5
1	Pattern description	0	0	1	1	4
2	Surveillance and outcome isolate counts (i.e. the ratios for each month)	0	1	0	0	5
3	Location summary	0	0	1	1	4
4	Pattern's history (displayed by clicking at the pattern ID)	1	3	2	0	0
5	Patient listing - source	0	0	1	2	2
6	Patient listing - location	0	0	2	1	3
7	Patient listing - antibiogram	0	0	2	2	2
8	Patient listing – specimen date	0	0	1	5	0
9	Patient listing – duplicate isolates' info	0	1	4	1	0
10	Cross pattern relationships	0	3	2	1	0
11	The "Show" button in the upper part of the screen (To show patterns with the selected attributes)	4	0	1	0	1
13	Group's importance score	1	1	3	0	1
14	Group's importance rank	0	1	3	1	1
15	Group's importance rating distribution	0	1	2	0	3
16	Group's reasons	2	3	0	1	0
17	Group's elaborated reasons	0	1	0	5	0
18	Are your ratings for 1-11 the same in all rounds?	0	0	0	0	6
19	Did you seriously select the clickable reasons	2	0	1	0	3

Post-study Survey Summary

		Strongly Disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly agree 5
1	The Infection Surveillance System website was easy to navigate.	0	0	0	4	2
2	I had no trouble reading the information displayed on screen.	0	0	0	3	3
3	The information was displayed in a way that was easy to understand.	0	0	2	3	1
4	The terminology was consistent with standard medical use	0	0	0	5	1
5	The instructions were clear and easy to understand.	0	0	0	5	1
6	The response time of the website was good.	0	0	0	4	2
7	It was useful to see the other panel member's votes and comments before I made my final decision.	0	1	0	2	3
8	Other group members contributed expertise that I found useful in my decision making	0	2	0	3	1
9	The process of using successive rounds of voting with the option of changing votes was a useful group decision- making process.	0	1	1	3	1
10	I would use such a group decision-making process again.	0	0	2	0	2
11	I would say the time required to complete the process each month was about right.	0	1	1	4	0
12	I made a sincere effort to review the data and to use the voting process as directed (please be honest).	0	0	0	5	1

		Very difficult	Difficult	About right	Easy	Very Easy
13	The difficulty involved in making judgments using the website information was:	0	2	3	1	0

Post Study Survey Summary (Continued)

Questions 14-18 are free text questions:

14. Describe your overall opinion of the use of the iterative group decision process to determine interpretations for the alerts.

    No response.

15. What are the strengths and the weaknesses of this process? What did you like and dislike?

    1 response: Gives one the opportunity to initially make a decision independently, but subsequently re-examine it in light of other's input. Good for learning.

16. How might the process be improved?

    1 response: I would suggest only two rounds, followed by results.

17. How might the website be improved?

    No response.

18. Other Comments (especially if you wish to elaborate on any of the questions in this survey)

    1 response: Don't stop now!

APPENDIX E

UAB IRB APPROVAL FORM

THE UNIVERSITY OF
ALABAMA AT BIRMINGHAM

Institutional Review Board for Human Use

Form 4: IRB Approval Form
Identification and Certification of Research
Projects Involving Human Subjects

The Institutional Review Board for Human Use (IRB) has an approved Multiple Project Assurance with the Department of Health and Human Services and is in compliance with 21 CFR Parts 50 and 56 and ICH GCP Guidelines. The Assurance became effective on January 1, 1999 and the approval period is for five years. The Assurance number is M-1149.

Principal Investigator:	WONG, DAISY Y.
Co-Investigator(s):	
Protocol Number:	**X000807003**
Protocol Title:	*Collaborative Interpreting for Knowledge Discovery*

The IRB reviewed and approved the above named project on 10-02-02. The review was conducted in accordance with UAB's Assurance of Compliance approved by the Department of Health and Human Services. This Project will be subject to Annual continuing review as provided in that Assurance.

This project received EXPEDITED review.

IRB Approval Date: 10-02-02

Date IRB Approval Issued: 10-02-02

Marilyn Doss

Marilyn Doss, M.A.
Vice Chair of the Institutional Review
Board for Human Use (IRB)

Investigators please note:

The IRB approved consent form used in the study must contain the IRB approval date and expiration date.

IRB approval is given for one year unless otherwise noted. For projects subject to annual review research activities may not continue past the one year anniversary of the IRB approval date.

Any modifications in the study methodology, protocol and/or consent form must be submitted for review and approval to the IRB prior to implementation.

Adverse Events and/or unanticipated risks to subjects or others at UAB or other participating institutions must be reported promptly to the IRB.

470 Administration Building
701 20th Street South
205.934.3789
Fax 205.934.1301
irb@uab.edu

The University of
Alabama at Birmingham
Mailing Address:
AB 470
1530 3RD AVE S
BIRMINGHAM AL 35294-0104

GRADUATE SCHOOL
UNIVERSITY OF ALABAMA AT BIRMINGHAM
DISSERTATION APPROVAL FORM
DOCTOR OF PHILOSOPHY

Name of Candidate Daisy Wong

Graduate Program Computer Science

Title of Dissertation Collaborative Interpreting for Knowledge Discovery

I certify that I have read this document and examined the student regarding its content. In my opinion, this dissertation conforms to acceptable standards of scholarly presentation and is adequate in scope and quality, and the attainments of this student are such that she may be recommended for the degree of Doctor of Philosophy.

Dissertation Committee:

Name	**Signature**
Warren T. Jones, **Chair**	Warren T. Jones
Stephen A. Moser	Stephen A. Moser
Barrett R. Bryant	Barrett R. Bryant
Robert M. Hyatt	Robert M. Hyatt
James Michael Hardin	J. Michael Hardin

Director of Graduate Program Warren T. Jones

Dean, UAB Graduate School ______

Date ______