IMPLEMENTATION OF A FUZZY RULE-BASED DECISION SUPPORT SYSTEM FOR THE IMMUNOHISTOCHEMICAL DIAGNOSIS OF SMALL B-CELL LYMPHOMAS

A Thesis presented to the Faculty of the Graduate School
University of Missouri

In Partial Fulfillment
Of the Requirements for the Degree

Master of Science

By

GERALD L. ARTHUR, M.D.

Yang Gong, MD, PhD, Thesis Supervisor

MAY 2009

The undersigned, appointed by the dean of the Graduate School, have examined this thesis entitled

IMPLEMENTATION OF A FUZZY RULE-BASED DECISION SUPPORT SYSTEM FOR THE IMMUNOHISTOCHEMICAL DIAGNOSIS OF SMALL B-CELL LYMPHOMAS

presented by Gerald Arthur,

a candidate for the degree of master of science,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Yang Gong

Professor Charles W. Caldwell

Professor Mihail Popescu

For Teddy, my wife,

whose idea this was,

and for all the good things that she

has brought into my life.

ACKNOWLEDGMENTS

I am appreciative of, and thankful for, the help of several people who have given me an abundance of assistance during the five year course of my studies for this Master's Degree in Health Informatics. Dr. Yang Gong, my advisor, provided me with a great introduction to Clinical Decision Support Systems in his HMI 8571 course which was a thoroughly enjoyable experience and an excellent foundation for thinking about such systems. Yang was also always willing to slice through the tangled and, occasionally impossible to accomplish, course requirements that were encountered in my unique academic pathway. Dr. Bill Caldwell has been a pathology colleague and friend for many years and, in addition, over the past five years he has been a remarkable mentor and guide for me in the unfamiliar worlds of academia, biomedical informatics and molecular biology. I am grateful for the many opportunities that he has arranged for me. Dr. Mihail Popescu has been a friend and teacher since my earliest experience with the University of Missouri during which time we both were National Library of Medicine Fellows in Biomedical Informatics. Then, separated by only a few cubicles, he introduced me to the intricacies and merits of MATLAB® and retrieved me safely from the computational black holes into which I found myself relentlessly drawn. As a professor, he taught two excellent courses in the domains of Knowledge Management and Ontology which were fascinating and, as I now appreciate, central to the whole idea of informatics.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....ii

LIST OF FIGURES.....v

LIST OF TABLES.....vi

LIST OF ABBREVIATIONS.....vii

ABSTRACT.....viii

Chapter

1. INTRODUCTION.....1

   Impact of the Genomic Revolution on Clinical Medicine

   The Challenge of Genomic Complexity

   The Need for Integration and Interpretation of Genomic Data in the Clinical Context

   Central Role of Surgical Pathology in Medicine

   The Cognitive Process in Surgical Pathology

2. THE IMPLEMENTATION OF CLINICAL DECISION SUPPORT SYSTEMS IN MEDICINE.....8

   Importance of Evidence-Based Medicine

   Justification for a Rule-Based Decision Support System

   The Advantages of a Fuzzy Rule-based Decision Support System

   Review of the Fuzzy Rule-Based Decision Support Literature

   Fuzzy Rule-Based Decision Support Systems in General Medicine

   Rule-Based Decision Support Systems in Surgical Pathology

3. MOTIVATION FOR DEVELOPING A PATHOLOGY DECISION SUPPORT SYSTEM.....17

Serve as a Vehicle for Establishing and Communicating National Best Practices in Pathology

To Reduce the Incidence of Inaccurate and Inconsistent Diagnoses

To Assist in the Identification and Reporting of Critical Clinically Relevant Pathologic Information

Significance and prevalence of antibody and nucleotide based probes in formulating a histopathological diagnosis

The challenge of integrating histopathology and pharmacogenomics

Receptivity of the pathology profession

4. DECRIPTION OF PILOT DEMONSTRATION PROJECT.....24

Design Objectives

Algorithm and Computational Platform

Construction of Rules

Function and Testing

5. Conclusions.....32

APPENDIX

1. FIGURES.....34

2. TABLES.....36

BIBLIOGRAPHY.....40

VITA.....49

LIST OF FIGURES

Figure Page

1. Triangular Membership Function Curves for Immunohistochemical Input Variables.....30

2. MATLAB® Fuzzy Logic Toolbox™ graphical user interface for viewing rules.....30

LIST OF TABLES

Table Page

1. Expected Immunophenotypes for Small B-Cell Lymphomas.....28
2. Rules for Features of Small B-Cell Lymphomas.....29
3. Positive Results for Extreme Input Values.....30
4. Positive Results for Marginal Input Values.....31

LIST OF ABBREVIATIONS

CGH	Comparative Genomic Hybridization
CLL	Chronic Lymphocytic Leukemia
DSS	Decision Support System
EBM	Evidence-Based Medicine
FCM	Fuzzy Cognitive Maps
FIS	Fuzzy Inference System
FL	Follicular Lymphoma
GUI	Graphical User Interface
HCL	Hairy Cell Leukemia
ICU	Intensive Care Unit
IHC	Immunohistochemical
LPL	Lymphoplasmacytic Lymphoma
MALT	Mucosa Associated Lymphoid Tissue Lymphoma
MCL	Mantle Cell Lymphoma
MZL	Marginal Zone Lymphoma
PDSS	Pathology Decision Support System
SNP	Single Nucleotide Polymorphism
TMD	Temperomandibular Disorder

IMPLEMENTATION OF A FUZZY RULE-BASED DECISION SUPPORT SYSTEM FOR THE IMMUNOHISTOCHEMICAL DIAGNOSIS OF SMALL B-CELL LYMPHOMAS

Gerald L Arthur

Dr. Yang Gong, Thesis Supervisor

ABSTRACT

Integrating the massive amount of new genomic information into medical practice will require new approaches to knowledge management. This need is particularly acute in surgical pathology where the evaluation of diagnostic tissue sections now requires assessment of the location and staining intensity for a wide range of proteins that are identified through the use of labeled antibodies. The possible combinations of visual features that may be observed threaten to exceed the limits of the human intellect.

Research was undertaken to construct a knowledge-based decision support system for this task. An algorithm written in MATLAB® and employing fuzzy logic was designed to interpret the staining results of eight antibodies that are used to classify lymphoid tumors. Testing with all possible combinations of antibody results (256) was consistent with the expected diagnoses. It appears that this method may be helpful in solving the information overload problem and may significantly improve the accuracy and consistency of disease diagnoses.

Introduction

The Impact of the Genomic Revolution on Clinical Medicine

The genomic revolution holds great promise for contributing to significant advances in clinical medicine by creating a new and more profound level of insight into human disease. The explosion of knowledge resulting from the application of the highly innovative experimental investigative techniques of molecular biology and genetics is evident from a number of parameters including the logarithmic growth of papers published in the field, sequences registered in Genbank or microarray results deposited in Gene Expression Omnibus. The domain of molecular biology and genetics produces an immense amount of data as a consequence of the development of new high-throughput analytic techniques such as second generation DNA sequencing, comparative genomic hybridization (CGH), and expression, single nucleotide polymorphism (SNP) and other types of microarrays. These remarkable technical advances have ignited a revolution in biomedical science that is only beginning to be incorporated into clinical medical care. The effective translation of the rapidly expanding knowledge of molecular biology into a new paradigm for medical practice probably represents the greatest and most revolutionary challenge to the medical profession since Virchow's elucidation of the cellular basis of disease one hundred and fifty years ago. Meeting this challenge will require health professionals to integrate a vast range

of new concepts and terminology into the cognitive analysis of disease processes. Pathological phenotypes will be best understood in reference to countless newly recognized, and yet to be discovered, molecular abnormalities. This new vision of health care is reflected in the idea of “personalized" or "genomic" medicine which envisions the determination of diagnosis, prognosis and design of a therapeutic plan that is tailored to the patient’s unique genetic makeup. It is clear that a new foundation is evolving for clinical medical knowledge and that it will be based on the principles of molecular biology rather than the concepts based on tissue and cellular abnormalities that have been employed in the past.

The Challenge of Genomic Data Complexity

Genomic, epigenomic, proteomic, transcriptomic and other types of molecular genetics data is voluminous and complicated. Much of the data generated in these fields can be converted into information meaningful to humans only with the assistance of computational software tools for knowledge management and discovery. There are currently a variety of bioinformatics techniques for interpreting this high-dimensional data that is provided through an array of computer software applications. However, the domain of clinical medicine has, in general, remained largely uninformed about these modalities and their potential utility in patient care in the era of genomic medicine. Thus, there is a critical need to stimulate the implementation of information and

knowledge management techniques that will support rational and intelligent data analysis in order to optimize the quality of clinical medical decisions.

The Need for Integration and Interpretation of Genomic Data in the Clinical Context

The integration of genomic data into health care requires that it be placed into the appropriate clinical context in relationship to the patient's phenotype. Ultimately, this may range from the definitive identification of a single pathognomonic genetic mutation to an indefinite, statistical interpretation of the significance of the hundreds of thousands of features identified through DNA sequencing or by expression, CGH, SNP or other microarray procedures. This type of data is already being generated by diagnostic studies in clinical medicine for the classification of multiple diseases, predominantly neoplastic and congenital disorders, and its use will likely grow exponentially. (Bagg, 2005; Baird et al., 2005; Cheang et al., 2008; Cook, Jr. & Scherer, 2008; Geschwind, 2003; Marshall et al., 2008; Smith et al., 2009)

Central Role of Surgical Pathology in Clinical Medicine

The histopathologic interpretation of multiple types of tissue and cytologic specimens provides crucial information the accuracy of which is critical for providing excellent health care. These tissue evaluations are generally viewed as "gold standard" medical facts that provide the highest quality, most reliable, diagnostic evidence available. There is likely no investigative modality that can

match the economical yield of information available to a skilled morphologist interpreting a tissue section stained only with basic hematoxylin and eosin. (Hicks et al., 2008; Rosai, 2007; Rosai, 2001) Nonetheless, pathologists have anticipated the genomic revolution by incorporating antibody-linked stains to accurately and sensitively identify the presence and location of multiple proteins in diagnostic tissue sections. This practice would appear to represent the first application of large-scale, effective proteomics in clinical medicine. To a somewhat lesser extent, pathologists have also integrated the use of nucleotide probes in the evaluation of DNA within tissue specimens. As genomic medicine evolves, pathologists will be required to vastly expand their skills in this arena which has been labeled morphogenomics and morphoproteomics. The realization of the ideal of personalized medicine will be dependent upon the recognition and evaluation of genomic and proteomic features that are unique to an individual and his disease. In order to meet this need, pathologists are now employing technologies that can identify complex molecular interactions at the tissue level. The highly objective and rigorous process of evaluating molecular abnormalities results in the ability to sub-classify patient and disease phenotypes permitting the determination of treatment plans specific for a particular disease and patient population. This knowledge is also invaluable in guiding the choice of additional algorithms and tools to be employed in the further diagnostic evaluation of the patient.

The Cognitive Process in Surgical Pathology

Diagnostic histopathology requires both visual and logical skills to accurately interpret microscopic images. The cognitive heuristics involved in the recognition of pathologic visual patterns are clearly related to training and experience but the decision-making process involved in this realm are poorly defined and will not be discussed in this paper. By contrast, the process of rationally analyzing and interpreting a set of objective facts can be readily elucidated and reasonably well modeled using various artificial intelligence tools such as decision trees. Thus, computer assistance may be expected to be usefully implemented to assist in the organization and interpretation of objective pathologic findings such as the incidence of various immunophenotypes or DNA abnormalities encountered in specific tumor types.

> "It is my impression that many of the cases encountered by pathologists require generation of several diagnostic possibilities (hypotheses), collection of and assignment of value to data, and use of these data to judge the likelihood that each hypothesis is either correct or incorrect. Much of this is rule-based behavior organized by diagnostic heuristics, and the success of the pathologist is dependent on his or her ability to understand how rules are applied to solving a scientific problem." (Foucar, 2001b)

However, in the process of implementing the new genomic technologies, pathologists are increasingly required to interpret and synthesize a rapidly expanding set of significant features to provide definitive disease classifications.

> "Although the adaptation of emerging technologies to clinical diagnostics may seem daunting in the information age, it really represents a continuation of the traditional role of the pathologist as the steward of translational application of new analytical discoveries to diagnostic medicine."(Finn, 2007e)

The number of items to be evaluated and correlated is rapidly approaching the limits of the human intellect and a geometric increase in complexity is expected to occur due to the continuing addition of myriad pertinent genomic and proteomic discoveries. This increase in knowledge has placed great demands upon the cognitive skills of even the most experienced and well-trained surgical pathologists. The number of facts to be considered in the interpretation of biopsies from even a single organ system has become overwhelming.

Transcending these cognitive limitations will certainly require computational support for the intelligent analysis of high-dimensional data. It is clear that the practice of pathology at the molecular level requires the use of translational bioinformatics tools. (Farkas, 2008; Finn, 2007d) Pathology spokesmen have been receptive to the need for integration of new bioinformatics knowledge management skills into the practice of surgical pathology and, in fact, view the use of these new technologies as an essential responsibility of pathologists.

> "Now, however, the practice of anatomical and clinical pathology is immersed in the same revolution impacting basic science: the ability to generate massive datasets in short time periods and

> unprecedented analytical capacity due to information and systems advances". (Finn, 2007c)

Bioinformatics software tools are not trivial to install and comprehend; however, many are open source and free for adaptation as desired. There are many bioinformatics applications that would be appropriate to consider implementing in pathology to illustrate and clarify important concepts in molecular biology that have relevance to the evaluation of tissue sections. Pertinent examples include Cytoscape for the depiction of biological pathways and their relationship to disease and GenePattern for multiple types of genomic analysis and graphical representation. It is likely that a working comprehension of genomic medicine cannot be realistically acquired without the aid of such computational assistance. All such tools may be considered to be decision support systems in that they provide a means of introducing important knowledge into the decision making process. However, this project is focused on the integration of another type of bioinformatics tool into histopathology, an expert rule-based system for assisting in the diagnostic process.

The Implementation of Clinical Decision Support Systems in Medicine

Importance of Evidence-Based Medicine

While new discoveries through research in clinical disciplines, molecular genetics and bioinformatics are an important source of advance in medical knowledge, it is also recognized that many of the problems in health care are not technical in nature but reflect organizational deficiencies in the field of knowledge management and the solutions to such problems require changes in policy more than in knowledge discovery and technology.(Patel et al., 2008b) Evidence-Based Medicine (EBM) is a concept that has been widely accepted in clinical medicine to improve the organization, dissemination and application of knowledge related to the best standards of health care. This term was first introduced into the medical literature by D.M. Eddy and he has continued to guide the evolution the concept which he has defined as:

> "a set of principles and methods intended to ensure that to the greatest extent possible, medical decisions, guidelines, and other types of policies are based on and consistent with good evidence of effectiveness and benefit." (Eddy, 2005)

Decision support systems represent one effective tool for validating, sharing, and applying knowledge resources in the medical community in keeping with the principles of EBM. This project intended to create a pilot DSS that

would demonstrate the feasibility of incorporating such a knowledge management tool into the daily practice of histopathology. The specific example selected to serve as a model diagnostic problem is the interpretation of immunohistochemical markers utilized in the evaluation of a specific class of lymphomas. The rules employed in this PDSS are based on a single paper on the diagnosis of lymphomas which was written by several experts in the field. (Higgins et al., 2008b)

Justification for a Rule-based Decision Support System

Clinical decision support systems (CDSS) may be classified as either knowledge-based or non-knowledge-based. Those in the former category attempt to enhance the clinician's innate cognitive ability to arrive at a correct diagnosis through the presentation of a set or rational and well-defined rules. By contrast, those in the latter category depend upon computer recognition of significant patterns of features using a variety of machine learning algorithms such as support vector machines or artificial neural networks. This approach essentially permits no human intervention other than in the choice of an appropriate population of medical data from which the machine "learns" the proper conclusions. The exact method by which the diagnosis is reached is extremely opaque and requires a degree of blind faith in the process to which a medical professional is unlikely to subscribe. It is difficult to imagine a physician or other responsible professional delegating his decision-making to a machine

with no knowledge of the method by which the decision was reached. By contrast, a rule-based system can document every single factor in the decision making process with references to papers by experts and, in addition, can provide ready access to knowledge resources containing extensive supplementary genomic information about sequences, expression, SNPs, cellular pathways, etc.

There are a number of other advantages to the use of a rule-based decision support system to implement Evidence Based Medicine. It provides a means of clearly communicating guidelines and protocols. Compared to the personal interpretation of a paper by a single practitioner whose interpretation may be erroneous or whose later recollection may be incomplete or inaccurate, the implementation of a computer based fuzzy logic system will ensure a uniform interpretation as recommended by experts in the field. Rule-based systems can be designed in a rational manner to achieve desired goals. As a matter of public health policy, it appears that any specific PDSS would represent the systematization and formalization of a consensus knowledge base derived from the accumulated experience of many domain experts and could be readily updated as advances in the field become recognized and accepted. At the same time, the decision-making process should support the possibility of favoring various types of desired outcomes as considered desirable for a given disease. For instance, one could create rules to favor various categories of diseases in the decision making strategies such as those that are most serious if left untreated

(prognostic approach), those that are most likely (probabilistic approach) or those which are most amenable to treatment (pragmatic approach). (Foucar, 2001a)

The Advantages of a Fuzzy Rule-Based Decision Support System

Medical decisions are based on inherently indistinct and poorly defined, that is to say "fuzzy", features and this is particularly evident in surgical pathology, one of the most “scientific” and objective of medical disciplines. It is rarely possible to be completely accurate in the description of a specific histological attribute. For instance, the average nuclear size in a given tumor may be judged to be "large" but invariably displays a range of sizes that in any specific case might be more realistically viewed as "sort of" large. That is, crisp or binary logic is imposed on a situation that is simply not that clearly defined. This is effectively recognized in the evaluation of immunohistochemical stains by generally employing a semi-quantitative scale for degrees of positive staining ranging from none to heavy (0 to 4+). In this situation it is expected that a range of staining will be encountered and that the final determination represents an accurate average of the stain results. It is clear that this and many other

evaluations in surgical pathology and medicine in general involve an element of subjectivity which is better accommodated by fuzzy rules.

A subset of rule-based systems utilizes "fuzzy" rather than "crisp" logic. Fuzzy logic is based on fuzzy set theory and recognizes that definitive boundaries are difficult to establish in continuous populations. Lofti Zadeh pioneered the concept of fuzzy logic in which membership in a particular class is a matter of degree rather than a binary either/or decision. (Zadeh, 1965; Zadeh, 1965) This approach is more reflective of the imprecision of real-life decisions in the clinical realm where the boundaries of parameters may be indistinct. (Licata 100-06;Ramesh et al. 334-38) As explained by Seising,

> "Apparently, it is very difficult to define sharp borders between various symptoms in the set of all symptoms and between various diseases in the set of diseases. Rather, we find smooth transitions from one entity to the other, and perhaps a very small variation might be the reason why a medical doctor diagnoses a patient with disease x instead of disease y." (Seising, 2006)

Review of the Fuzzy Rule-Based Decision Support Literature

The medical literature describing rule-based decision support systems was reviewed with particular emphasis on those employed in the field of surgical pathology and that were using fuzzy logic. The results of the review confirm that relatively few computer assisted decision support systems of any type have been

implemented for pathology. The majority of these deal with the interpretation of clinical pathology tests or to decisions pertaining to the ordering of the appropriate test for the clinical situation. The latter are usually invoked at the time of test-ordering as computer assisted physician order entry modules. Several papers were identified in the general medical literature and are included because they provide good illustrations of the fuzzy logic methodology.

Fuzzy rule-based decision support systems in general medicine

Bates and Young (Bates & Young, 2003) applied fuzzy logic to a simplified model of decision support for the administration of intra-venous fluids in the intensive care unit setting (ICU). They related five fuzzy sets of rates of intravenous fluid administration (IFR) to three sets each of the determining factors of mean arterial pressure (MAP) and hourly urine output (HUO). The result is a transformation of two input variables categorized as low, normal or high into five action states classified as rates of infusion of low, maintenance, moderate, high or very high.

Binaghi, Gallo, Ghiselli, Levrini and Biondi (Binaghi et al., 2008) described a fuzzy rule-based system designed to support diagnostic reasoning about temperomandibular disorders (TMD). This was integrated with a web-based source of protocols presenting a standardized methodology for the classification of TMD that had been determined by expert clinicians. The authors

found that the most critical design consideration was the assignment of knowledge elements to appropriate membership functions. Ultimately, they chose a simple triangular membership curve and the max-min deductive inference algorithm. Performance testing was performed on 50 cases ranging from 18 to 60 years of age and equally divided as to gender. The results were consistent with the clinical diagnosis rendered by experts in 100% of cases.

Rule-Based Decision Support Systems in Surgical Pathology

Foran, Comaniciu, Meer and Goodell (Foran et al., 2000) describe a PDSS consisting of two components: a distributed telemicroscopy system and an intelligent image repository. This system was designed to assist medical technologists in obtaining interactive consultation through a robotic microscope controlled remotely and to automatically access a database for retrieval of images and diagnoses. They primarily addressed the morphologic evaluation of histopathologic images by mathematical descriptors of cellular characteristics such as nuclear size and shape, cell size, cytoplasm texture and color that enabled the retrieval of comparable pathologic images from a data repository for comparison by diagnosticians. A similarity measure was devised that incorporated a summation of normalized distances from the center of each diagnostic class feature. They were able to demonstrate a correct classification rate of 83% when classifying cells into four groups, chronic lymphocytic leukemia, follicular center cell lymphoma, mantle cell lymphoma and normal.

Thus, the decision algorithm is based on the computational determination of features and is entirely machine determined.

Asare, Ellis and Caldwell developed a decision support system for cell immunophenotyping by flow cytometry. (Asare et al., 2002) This system was designed primarily as a quality improvement mechanism. They utilized various quality assurance mechanisms including delta checks and sums of monoclonal antibody fluorescence. In addition, results were presented to clinicians in graphic formats illustrating results over time and this was found to improve interpretability. The authors also explored the possibility of integrating an automated check for the identification of a change in immunoglobulin light chain expression between subsequent specimens from the same patient. This switch is unlikely to occur in the usual course of the disease and, if detected, would provide a prompt for the need to reexamine the findings. A decrease in clerical errors using this system was observed. The authors were able to document a decrease in the number of out-of-range results from 58% to 2% over a three year period.

Another decision support system for the interpretation of flow cytometry immunophenotyping data was reported by Nguyen et al. (Nguyen et al., 2008) This was constructed using a database containing the incidence of antigens expressed by various hematologic neoplasms as determined from scientific journal articles. The system analyzed 44 antigen markers reported as either positive or negative which were used to characterize 37 hematologic neoplasms.

A mathematical algorithm was derived to compare the observed immunophenotypes with those reported in the literature in which an antigen for a particular tumor was considered to be characteristically positive if more than 50% of cases were reported to be positive based on literature review. This DSS is designed to be informative to pathologists in training by providing a relatively broad differential diagnosis for an immunophenotype. This is an appropriate goal since, as noted by the authors, a definitive diagnosis of these neoplasms cannot be achieved without additional histologic confirmation. This system is based on a Bayesian decision making approach based on population sampling rather than representing an expert rule-based algorithm.

A decision support system that combines a fuzzy rule-based logic integrated with a machine learning component was reported by Papageorgiou et al. (Papageorgiou et al., 2006) They developed a method for diagnosing the grade of urinary bladder tumors using fuzzy cognitive maps (FCM) with an unsupervised artificial neural network learning algorithm. They defined eight histopathologic cellular features that are commonly used to judge tumor grade. Fuzzy cognitive maps are a soft computing methodology that combine fuzzy logic and neuro-computational techniques. Histopathological concepts are represented as nodes and their interactions as weighted edges. These are then formed into models according to dynamics determined by experts in the field. Their fuzzy logic model employed seven triangular membership functions for

specific histologic features. The diagnoses were inferred using individual rules weighted as valued by linguistic variables ranging from “very, very low” to “very, very high”. The opinion of experts was formulated as a weight matrix for the tumor grading model. The authors then submitted this matrix to modification by an unsupervised learning algorithm in order to ensure that the grade of the tumor as determined by the artificial neural network successfully matched that of the experts. Their system achieved an accuracy of 72.5%, 74.4% and 95.6% for tumors or grade I, II and III respectively as compared with expert evaluation. While, as the authors claim, this methodology provides a degree of transparency as to the rules on which the classification is based, the ultimate diagnostic rules are machine determined and are thus quite opaque.

Motivation for Developing a Pathology Decision Support System

To Serve as a Vehicle for Establishing and Communicating National Best Practices in Pathology

The construction of rule-based expert systems would seem to follow the best precepts of evidence based medicine (EBM). Wick, Bourne, Patterson, and Mills view EBM as being built on several elements: (1) the replacement of the traditional authoritative paradigm with an analytical approach; (2) the analysis of information using quantitative methods; (3) the integration of the best evidence

into protocols or guidelines that affect clinical practice; and (4) a continuous cycle of evaluation and improvement of these guidelines. (Wick et al., 2005)

There may be disagreement on exactly the best manner of obtaining and interpreting the data utilized for constructing the rules; however, it is assumed that organized pathology would assume the responsibility for assembling a body of recognized experts to oversee this process. In fact, leaders in the field of pathology have already begun to develop EBM guidelines for the comprehensive reporting of the clinically essential findings in cancer specimens and this very organized project may be expected to yield many additional useful algorithms and protocols for pathologists. (2009b)

To Reduce the Incidence of Inaccurate and Inconsistent Diagnoses

Developing the ability to comprehend, evaluate and integrate genomic information with features of a disease phenotype is a challenge facing all clinicians, but is arguably most imperative for surgical pathologists. The rapidly increasing knowledge of disease mechanisms to be recognized at the molecular level has placed great demands upon the integrative intellectual skills of even the most experienced, well-trained, and specialized surgical pathologists. It has long been recognized that there is a significant lack of consistency in the diagnostic evaluation of biopsy material even among expert surgical pathologists. (Clarke et al., 2004; Elsheikh et al., 2008) The consensus estimate of the incidence of

inaccurate diagnoses in surgical pathology is about 0.5% to 1.0%. (Nakhleh, 2006; Nakhleh, 2008) As a result of the increasingly complex knowledge domain for which pathologists are now responsible, it should be anticipated that this loss of diagnostic consistency is likely to increase dramatically. Unless corrected, these errors in surgical pathology diagnoses will have a great negative impact on the goal of achieving personalized medicine since many of the promised benefits of the genomic revolution depend upon accurate and reproducible diagnoses.

To Assist in the Identification and Reporting of Critical Clinically Relevant Pathologic Information

Recognition of the rapidly expanding catalogue of diseases newly defined in terms of molecular biology demands that pathologists incorporate information from a number of genomic knowledge bases into their diagnostic reasoning. This supplemental knowledge can enlarge their areas of competence by presenting organized information about the proteins that have been shown to be significant in the pathogenesis of neoplasia and other diseases. With the increasing integration of molecular genetics knowledge and clinical medicine, pathologists will be expected to identify specific features of tissue protein expression and to explain their diagnostic, prognostic and therapeutic significance.(Costa, 2008; Salto-Tellez, 2007) This extraction of knowledge from a complex mass of disparate features may be enhanced by computational support.

An extensive corpus of fundamental medical knowledge has been gained through decades of the scientific examination of tissues stained with the basic hematoxylin and eosin stain. Countless studies have identified ever more precise diagnostic entities based on fine visual and logical distinctions evident from alterations of cellular morphology. However, over the past several decades, this process has accelerated due to the application of antibody-linked stains (immunohistochemistry - IHC) which permits the identification of many cellular proteins that play an essential role in the molecular pathology of a wide range of diseases. A similar technology using in situ hybridization of fluorescent oligonucleotide probes (FISH) has resulted in the ability to interrogate DNA for mutations, deletions and duplications. These diagnostic tools, referred to as morphoproteomics and morphogenomics, have permitted pathologists to implement the most comprehensive functional application of genomics and proteomics in clinical medicine. As a consequence of this transformative technology that provides the ability to perceive sub-cellular pathogenic mechanisms, great demands have been imposed upon the integrative intellectual skills of even the most experienced and well-trained surgical pathologists. The mental retention and recall of the number of facts to be considered in the interpretation of biopsies from even a single organ system has now become problematic.

Significance and Prevalence of Antibody and Nucleotide-Based Probes in Reaching a Histopathologic Diagnosis

A large number of papers have demonstrated that the sub-classification of various neoplasms and other diseases on the basis of gene expression signatures better reflects the biology of the disease and provides stratification for the design of improved therapeutic regimens. (Baird et al., 2005; Cheang et al., 2008; Cheng et al., 2009; Goldstein et al., 2008; Marchionni et al., 2008; Maruya et al., 2004; Ramaswamy et al., 2003; Reis-Filho et al., 2006; Weigelt et al., 2008) Similarly, discrimination among tumor types may be accomplished through observation of antibody or nucleotide probe reactivity in microscopic sections. The use of this diagnostic modality is anticipated to increase exponentially over the coming decades and will result in a complexity of information processing that will be difficult to integrate into the diagnostic process without computer assistance.

There are now a number of commercial genomic assays utilized in the practice of oncology. These include MammaPrint, OncotypeDx, Theros and MapQuantDx for the prediction of clinical prognosis in breast cancer (Sotiriou & Pusztai, 2009) and Tissue of Origin (Pathwork Diagnostics (2009d)), CancerType ID® (bioTheranostics(2009a)) and MiRview™ (Rosetta Genomics(2009e)) for evaluating metastatic cancer of unknown primary site. It is anticipated that multi-gene assays of this type will be routinely requested by clinicians and the results will be increasingly determined from formalin-fixed, paraffin-embedded tissues that are processed in the surgical pathology laboratory. As discussed

above, morphologic evaluation is also required for the design of a customized therapeutic regimen. The significance of these clinically important facts will be catalogued, organized and documented within the PDSS to ensure that the final report reflects their evaluation.

Recognition of the rapidly expanding catalogue of diseases newly defined in terms of molecular biology demands that pathologists incorporate information from a number of genomic knowledge bases into their diagnostic reasoning. This supplemental knowledge can enlarge their areas of competence by presenting organized information about proteins that have been shown to be significant in the pathogenesis of neoplasia and other diseases.

The Challenge of Integrating Histopathology and Pharmacogenomics

A primary requirement for the delivery of personalized medical care is the ability to identify therapeutic agents that are optimal for a particular patient, the specific sub-type of disease with which they are afflicted and the drug efficacy and toxicity expected from the patient's unique genetics . The immunohistochemical identification of numerous proteins predicting the effectiveness of therapeutic agents has already become the standard for determining indications for their administration. The most well known of these pharmaceutical agents include tamoxifen, trastuzumab (Herceptin), imatinib mesylate (Gleevec), and gefitinib (Iressa). The surgical pathologist has the ability to determine both the presence

and micro anatomic location of the target protein and this combination provides a great deal of useful information in regard to state of activation, differential involvement of cell populations and potential molecular pathways affected by the drug. (Brown, 2009; Hunt, 2009) In view of these capabilities, it should be anticipated that surgical pathology will be the domain with primary responsibility for driving the realization of personalized medicine. (Hess, 2002) However, this effort must be supported by effective knowledge management systems and these could be included within the framework of a PDSS.

Receptivity of the Pathology Profession

Leaders in the field of pathology have been receptive to the integration of new molecular genetics and informatics skills into the practice of surgical pathology and, in fact, view the use of these new technologies as an essential responsibility of pathologists.(Costa, 2008; Finn, 2007b; Friedman, 2008; Levenson, 2008) Furthermore, the importance of genomics and bioinformatics in the new schema of intelligent data analysis is well appreciated.

> "As the pace of technological advancement accelerates, it behooves us to harness the potential of emerging informatics, genomic, and proteomic applications for the optimization of our specialty."(Finn, 2007a)

Therefore, there appears to be general agreement that something needs to be done to assure the quality of surgical pathology diagnoses. "Pathologists

reluctant to look at decision-making by examining its smallest, simplest parts should offer an alternative approach to improving diagnosis." (Foucar, 2001c)

Description of Pilot Demonstration Project

Design Objectives

Some definitions of a CDSS include the presence of inherent reasoning by an inference engine as an essential feature. (Berner & La Lande, 2007; Greenes, 2007) Others are broader and emphasize the concepts of knowledge management that may provide pertinent, structured information at critical stages of clinical evaluation to enhance the effectiveness of the decision-making process. (Cimino, 2007; Glaser & Hongsermeier, 2007; Hongsermeier et al., 2007) This view of CDSS holds that they are "computer systems designed to impact clinical decision making about individual patients at the point in time that these decisions are made." (Berner & La Lande, 2007)

Although this PDSS does contain an inference engine, it is not intended to replace the diagnostic skills of a pathologist. Rather, the system is intended to enable pathologists to be able to consult an expert, rule-based clinical decision support system to assure that all significant facts revealed by immunohistochemical stains have been considered and no pathognomonic features have been overlooked in the diagnostic process. This tool will also help

assure that in the presence of a specific immunophenotype, the diagnosis is one that a reasonable pathologist would have made thus helping to prevent errant diagnoses due to misapprehension of an obvious diagnostic element. This PDSS will also function as a method of knowledge organization and representation that creates a system for tracking the many specific detailed facts that should be considered in the course of fashioning a diagnosis. Thus it may be viewed as providing a bookkeeping function that works in the background to evaluate the consistency of objective evidence and will alert the diagnostician to a failure to consider relevant facts. This should provide a substantial aid in eliminating the errors of omission that have been viewed as a prominent cause of medical errors. (Patel et al., 2008a)

This project specifically intended to design a PDSS that does not incorporate any form of supervised or unsupervised machine learning such as artificial neural networks, support vector machines or other forms of artificial intelligence associated with pattern recognition computation. This principle is held due to the author's personal experience, and close long-term observations of other pathologists, in the process of formulating histopathologic diagnoses. From this viewpoint it appears evident that diagnosticians must have rational and documentable reasons for arriving at a definitive conclusion. To date, the forms of artificial intelligence alluded to above are not capable of providing this information.

The PDSS design should be able to accommodate the continuing expansion of the domain of morphoproteomics. This will provide essential support permitting pathologists to enlarge their areas of diagnostic expertise in an organized manner to include the identification of an ever-increasing number of proteins that have been shown to be critical to the pathogenesis of neoplasia as well as other diseases.

This pilot project is intended to construct and evaluate the use of a fuzzy inference system (FIS) as the basic central module for a decision support system. After demonstrating its utility in interpreting a reasonably complex set of diagnostic features, it may then be applied to a wide range of essentially similar diagnostic settings in which the basic function is unchanged. Therefore, it is anticipated that larger diagnostic problems may be addressed by employing several or many of these FIS modules in parallel to evaluate diseases which display coherent and well-defined subsets of morphoproteomic features. Control of the flow to these modules could be implemented either with additional fuzzy logic systems or by means of decision trees implementing simple if-then statements. User-friendliness is always an important consideration in the design of decision support systems. Accomplishing this goal appears to entail the implementation of a GUI and that is planned to be accomplished in the near future. Ultimately, it is envisioned that the entire PDSS should be established in a Windows® operating system in order to be available to the widest audience.

The PDSS should be capable of accurately reflecting the nuances of expert opinions. The fuzzy rule-based approach can support considerable fine-tuning of the diagnostic process in accordance with expert opinion. At the same time, this is accomplished in a rigorously reliable manner that may be readily transmitted to practitioners as a computational tool. As discussed above, fuzzy-logic membership functions may be portrayed graphically as a variety of lines or curves representing the degree to which a feature is present in a particular instance. By altering these shapes, their weighting and the degree to which they overlap, one may define quite precisely, the contribution of a particular feature to the final inferred decision. In terms of pathology, this would allow for a consensus opinion about the relative importance of specific features in the estimation of experts to be relatively realistically and accurately imposed on a set of observations.

While this PDSS may be based on rules objectively defined by the consensus opinions of experts, it may also be customized to reflect interpretations based on personal experience or alternative, dissenting opinions. This flexibility will be important if the PDSS is to become a reliable tool for the individual pathologist.

Algorithm and Computational Platform

The pilot PDSS is written using the high-level programming language MATLAB® which is a software application that integrates computation,

visualization and programming functions. MATLAB® provides a Fuzzy Logic Toolbox™ which contains a Graphical User Interface (GUI) for the development of fuzzy logic programs. MATLAB® supports the Microsoft Component Object Model and therefore programs developed with it may be ported to Windows operating systems.

Fuzzy logic supplies a means of mapping a set of input variables to another set of output variables using antecedent and consequent (if-then) rules that are tolerant of imprecision and reflect the qualitative descriptive terms of natural language. Thus there are four components to the implementation of the fuzzy logic process: (1) delineating the terms and variability contained within the rules; (2) specifying an algorithm for interpreting the input values and assigning appropriate values to the output space; (3) aggregating the output of all rules employed and; (4) "defuzzification" of the aggregate output to assign a single output value. The input and output values are transformed into fuzzy values by assignment to curves defining the range of possible values of the variables. These curves are called "membership functions". The membership functions, aggregation and defuzzification can be interpreted using a superset of standard Boolean logic that can account for set memberships that are continuous rather than binary. (2009c)

Construction of Rules

It is accepted that the definitive diagnosis of lymphomas requires the interpretation of a panel of antibodies to the multiple significant proteins that may be expressed by a particular tumor. Therefore, this PDSS is designed specifically for implementation in association with the demonstration of tissue proteins by antibody-based stains. A group of experts has recently published an article based on a review of the literature as well as their personal experience that specifies detailed criteria for the diagnosis of hematolymphoid malignancies. (Higgins et al., 2008a) The rules employed in the decision support system were extracted from the recommendations of these authors. They are largely based on a table of the expected immunophenotypic findings as defined by eight antigens for classifying seven neoplasms within the small B-cell lymphoma group. (Table 1) The fuzzy inference system is based on the input of a pathologist's evaluation of the intensity of staining for each of these antibodies on a value scale ranging from 1 to 5 corresponding to the standard semi-quantitative grading scheme.

Twenty-four (24) rules were devised that captured the diagnostic discriminations between tumor types that display differential degrees of staining. (Table 2) The rules were interpreted using the superset of Boolean logic in which AND and OR are represented by the minimum and maximum of the membership functions respectively. Membership functions for both the input and output variables were defined using three overlapping triangular curves. (Figure 1)

Truncation or "implication" and aggregation of the output values were based on the minimum or the maximum respectively. Defuzzification was performed using the largest value of the maximum aggregated result. The Matlab® Fuzzy Logic Toolbox™ provides a Rule Viewer GUI that simplifies the visualization and interpretation of the entire fuzzy inference system. (Figure 2)

Function and Testing

At this point in the evolution of the PDSS, the aim is to confirm that a crisp output may be obtained from crisp inputs in order to confirm that the algorithm is capable of providing accurate answers in clear cut situations. The algorithm has been tested using a number of matrices as input. The Matlab® Fuzzy Logic Toolbox™ provides a function for automatically testing input values using a matrix format. For this purpose, a matrix containing all 256 (2^8) possible input values limited to either 0 or 5 (clearly negative or strongly positive staining respectively) was constructed. Only 21 of the possible 256 combinations of antigen staining were diagnostic and, of these, 20 were diagnostic of a single tumor. (Table 3) All diagnoses were consistent with the expert-determined immunophenotype that should be observed for the specific tumor. One tumor, MALT, was identified by only one set of antigens which is to be expected as it expresses only a single antigen out of the set of eight. Other tumors were diagnosed more than once, which is explained by the presence of one or more proteins that may or may not be expressed in those tumors. Therefore, these

tumors have several options for diagnostic antibody patterns. One pattern of antigen features yielded diagnoses of both MCL and CLL which is correct since these two tumors cannot be differentiated by that particular pattern of antigen expression. To be specific, no false positive or negative diagnoses were identified. Therefore, it is concluded that the algorithm has faithfully reflected the opinion of the expert hematopoietic pathologists in all instances and may be relied upon as a decision support system. A matrix containing the less clear-cut input values of 1 or 4 yielded similar findings, but with two additional diagnoses of MZL and one of HCL. (Table 4) In additional testing, value vectors in the mid-range (2, 3) resulted in indeterminate results of 0.5 in all cases.

Research is underway to evaluate the variation of results with alterations in membership functions and parameter weighting. This will provide a practical understanding for designing implementations that will provide the desired variability in outcomes related to the importance of particular features as would be established by experts.

Conclusions

The need for decision support systems in clinical medicine is clearly evident from a variety of observations. The genomic revolution has had a significant impact on the medical domain already and it is obviously increasing at a rapid rate. This new knowledge needs to be integrated with that of classical medicine in a functional manner. However, the volume of information is overwhelming and cannot be fully utilized without computational support. Fortunately the fields of biomedical informatics, computer science and knowledge management and engineering have provided tools with which to address this problem. Among the potential solutions are decision support systems of various types.

Two additional conclusions are also derived from a review of the literature: (1) there are few decision support systems designed for surgical pathology and those that have been identified utilize some type of machine learning methodology; (2) a large number of recognized leaders in the field of surgical pathology have noted the challenges of the new era and have advocated the

acceptance of decision support tools. This paper argues for the implementation of a rule-based fuzzy inference system as the optimal application for the domain of surgical pathology for reasons related to the uncertainty inherent in medical practice and the transparency of the method.

This project has accomplished the aim of constructing a pilot fuzzy rule-based model that, after systematic testing, appears to function with a high degree of accuracy that should be entirely adequate for its intended purpose in the clinical realm. Compared to previous reports of fuzzy rule-based systems, this model has a larger number of input and output values. Therefore, this pilot model has demonstrated that a fuzzy inference system is a practical application for the interpretation of the complex decision processes encountered in pathology and clinical medicine in general. This system is capable of not only supporting the diagnostic process but also serving as a means of sharing the best practices of Evidence-Based Medicine through the incorporation of well structured protocols. Furthermore, additional design elements have been proposed that are likely to enhance its utilization by surgical pathologists.

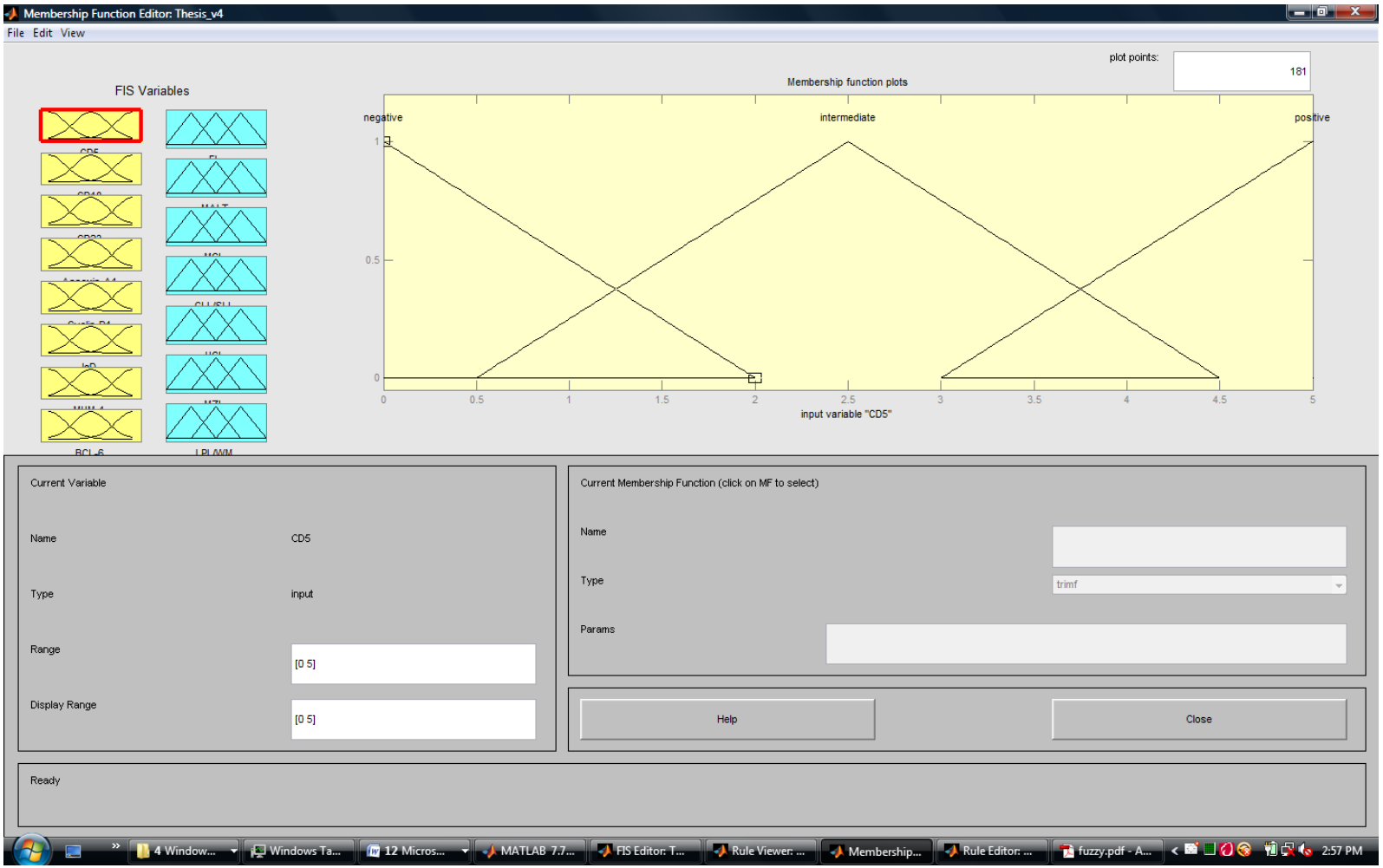

Figure 1. Triangular membership function curves for immunohistochemical input variables.

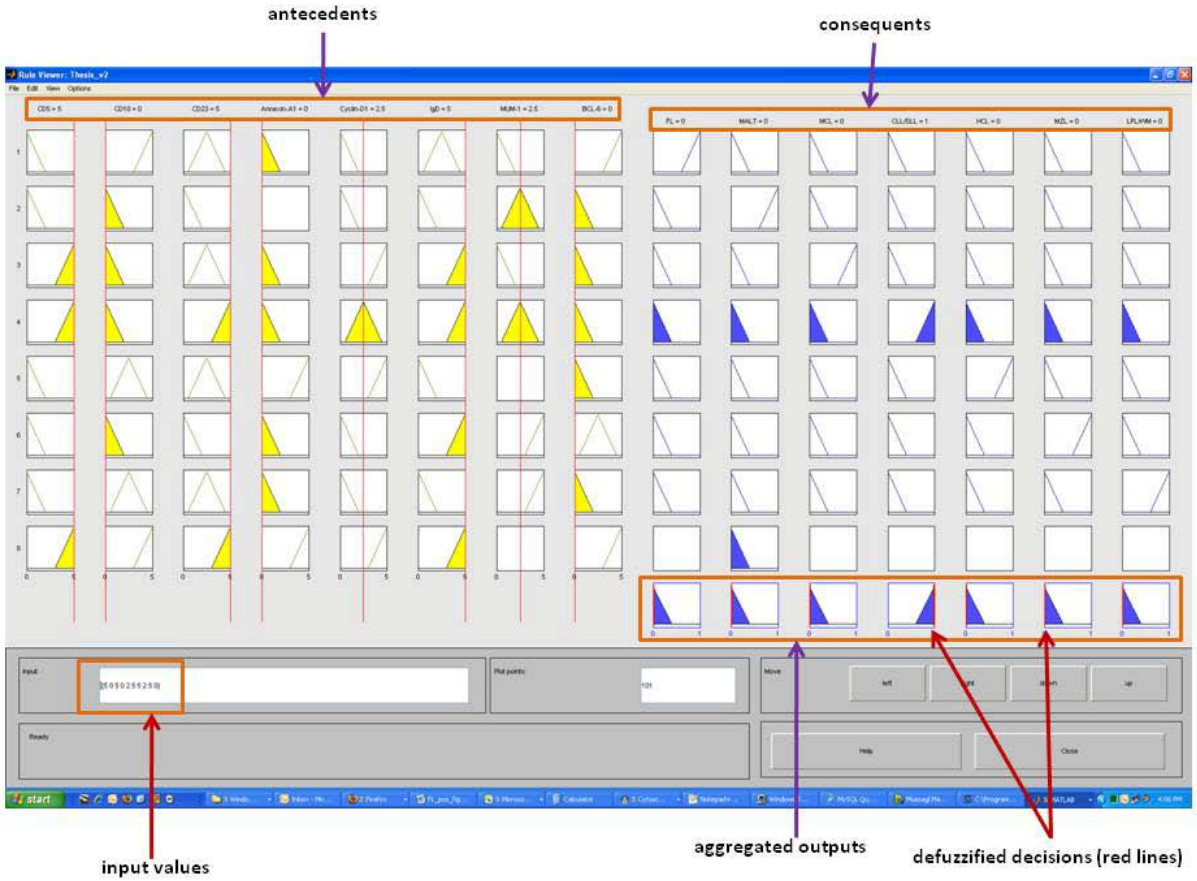

Figure 2. MATLAB® Fuzzy Logic Toolbox™ graphical user interface for viewing rules.

Neoplasm	CD5	CD10	CD23	Annexin-A1	Cyclin D-1	IgD	MUM-1	BCL-6
FL	-	+	-/+	-	-	-/+	-	+
MALT	-	-	-	NT	-	-	+/-	-
MCL	+	-	-/+	-	+	+	-	-
CLL/SLL	+	-	+	-	-/+	+	+/-	-
HCL	-	-/+	+	+	+	-	NT	-
MZL	-	-	-	-	-	+/-	+	-/+
LPL/WM	-	-/+	-/+	-	-	-	+	-

Table 1. Table of antibody staining responses to eight proteins demonstrated by seven Small B-Cell Lymphomas. (Higgins, 2008)

IF	CD5	&	CD10	&	CD23	&	Annex-A1	&	Cyc D-1	&	IgD	&	MUM-1	&	BCL-6	=	FL	MALT	MCL	CLL	HCL	MZL	LPL
	-		+				-		-				-		+		+	-	-	-	-	-	-
	-		-		-				-		-		+		-		-	+	-	-	-	-	-
	+		-				-		+		+		-		-		-	-	+	-	-	-	-
	+		-		+		-				+				-		-	-	-	+	-	-	-
	-						+		+						-		-	-	-	-	+	-	-
	-		-		-		-		-				+				-	-	-	-	-	+	-
	-						-		-		-		+		-		-	-	-	-	-	-	+

IF	CD5	OR	CD10	OR	CD23	OR	Annex-A1	OR	Cyc D-1	OR	IgD	OR	MUM-1	OR	BCL-6	=	FL	MALT	MCL	CLL	HCL	MZL	LPL
	+		+		+		+		+		+				+			-					
			-												-		-						
													-					-					-
			+				+		-		-		+						-				-
	-		+		-		+				-				+					-			-
	+						+		-		+				+						-		-
	+		+		+		+		+				-									-	
							+										-		-	-		-	-
															+			-	-	-	-		-
			+															-	-	-		-	
													+				-		-		-		

Table 2. Rules for interpreting immunohistochemical features for small B-cell lymphomas derived from opinions of expert hematopathologists. These are the rules as implemented with the MATLAB® Fuzzy Logic Toolbox™.

CD5	CD10	CD23	Annex -A1	Cyclin D-1	IgD	MUM-1	BCL-6	FL	MALT	MCL	CLL	HCL	MZL	LPW	DX
0	0	0	0	0	0	5	0	0	1	0	0	0	0	0	MALT
0	0	0	0	0	5	5	0	0	0	0	0	0	1	0	MZL
0	0	0	0	0	5	5	5	0	0	0	0	0	1	0	MZL
0	0	0	5	5	0	0	0	0	0	0	0	1	0	0	HCL
0	0	0	5	5	0	5	0	0	0	0	0	1	0	0	HCL
0	0	5	0	0	0	5	0	0	0	0	0	0	0	1	LPL
0	0	5	5	5	0	0	0	0	0	0	0	1	0	0	HCL
0	5	0	0	0	0	0	5	1	0	0	0	0	0	0	FL
0	5	0	0	0	5	0	5	1	0	0	0	0	0	0	FL
0	5	0	5	5	0	0	0	0	0	0	0	1	0	0	HCL
0	5	0	5	5	0	5	0	0	0	0	0	1	0	0	HCL
0	5	5	0	0	0	0	5	1	0	0	0	0	0	0	FL
0	5	5	0	0	0	5	0	0	0	0	0	0	0	1	LPL
0	5	5	0	0	5	0	5	1	0	0	0	0	0	0	FL
0	5	5	5	5	0	0	0	0	0	0	0	1	0	0	HCL
0	5	5	5	5	0	5	0	0	0	0	0	1	0	0	HCL
5	0	0	0	5	5	0	0	0	0	1	0	0	0	0	MCL
5	0	5	0	0	5	0	0	0	0	0	1	0	0	0	CLL
5	0	5	0	0	5	5	0	0	0	0	1	0	0	0	CLL
5	0	5	0	5	5	0	0	0	0	1	1	0	0	0	MCL/CLL
5	0	5	0	5	5	5	0	0	0	0	1	0	0	0	CLL

Table 3. Positive test results for all possible combinations (256) of input values limited to extremes (0,5) of the scale. The first eight terms represent names of antigens, the next seven terms are potential lymphoma diagnoses (output values) and the final column is the diagnosis. Output value of zero is negative and one is positive.

CD5	CD10	CD23	Annex -A1	Cyclin D-1	IgD	MUM-1	BCL-6	FL	MALT	MCL	CLL	HCL	MZL	LPW	DX
1	1	1	1	1	1	4	1	0.2	1	0.2	0.19	0.2	1	0.19	MALT/MZL
1	1	1	1	1	1	4	4	0.2	0.2	0.2	0.19	0.2	1	0.19	MZL
1	1	1	1	1	4	4	1	0.2	0.2	0.2	0.19	0.2	1	0.19	MZL
1	1	1	1	1	4	4	4	0.2	0.2	0.2	0.19	0.2	1	0.19	MZL
1	1	1	4	1	1	4	1	0.2	1	0.2	0.19	0.2	0.2	0.19	MALT
1	1	1	4	4	1	1	1	0.2	0.2	0.2	0.19	1	0.2	0.19	HCL
1	1	1	4	4	1	4	1	0.2	0.2	0.2	0.19	1	0.2	0.19	HCL
1	1	4	1	1	1	4	1	0.2	0.2	0.2	0.2	0.2	0.2	1	LPL
1	1	4	4	4	1	1	1	0.2	0.2	0.2	0.2	1	0.2	0.2	HCL
1	1	4	4	4	1	4	1	0.2	0.2	0.2	0.2	1	0.2	0.2	HCL
1	4	1	1	1	1	1	4	1	0.2	0.2	0.19	0.2	0.2	0.19	FL
1	4	1	1	1	4	1	4	1	0.2	0.2	0.19	0.2	0.2	0.19	FL
1	4	1	4	4	1	1	1	0.2	0.2	0.2	0.19	1	0.2	0.19	HCL
1	4	1	4	4	1	4	1	0.2	0.2	0.2	0.19	1	0.2	0.19	HCL
1	4	4	1	1	1	1	4	1	0.2	0.2	0.2	0.2	0.2	0.2	FL
1	4	4	1	1	1	4	1	0.2	0.2	0.2	0.2	0.2	0.2	1	LPL
1	4	4	1	1	4	1	4	1	0.2	0.2	0.2	0.2	0.2	0.2	FL
1	4	4	4	4	1	1	1	0.2	0.2	0.2	0.2	1	0.2	0.2	HCL
1	4	4	4	4	1	4	1	0.2	0.2	0.2	0.2	1	0.2	0.2	HCL
4	1	1	1	4	4	1	1	0.2	0.2	1	0.19	0.2	0.2	0.19	MCL
4	1	4	1	1	4	1	1	0.2	0.2	0.2	1	0.2	0.2	0.2	CLL
4	1	4	1	1	4	4	1	0.2	0.2	0.2	1	0.2	0.2	0.2	CLL
4	1	4	1	4	4	1	1	0.2	0.2	1	1	0.2	0.2	0.2	MCL/CLL
4	1	4	1	4	4	4	1	0.2	0.2	0.2	1	0.2	0.2	0.2	CLL

Table 4. Positive test results for all possible combinations (256) of input values limited to less definitive values (1, 4) of the scale. The first eight terms represent names of antigens, the next seven terms are potential lymphoma diagnoses (output values) and the final column is the diagnosis. Output value of zero is negative and one is positive. Others represent relative diagnostic probability.

Reference List

BioTheranostics (2009a). BioTheranostics [On-line]. Available: http://www.aviaradx.com/

CAP-Synoptic Reporting (2009b). College of American Pathologists [On-line]. Available: http://www.cap.org/apps/cap.portal?_nfpb=true&cntvwrPtlt_actionOverride=%2Fportlets%2FcontentViewer%2Fshow&_windowLabel=cntvwrPtlt&cntvwrPtlt{actionForm.contentReference}=committees%2Fcancer%2Fcancer_protocols%2Fprotocols_index.html&_state=maximized&_pageLabel=cntvwr

Matlab Fuzzy Logic Toolbox (2009c). Mathworks Inc. [On-line]. Available: http://www.mathworks.com/access/helpdesk/help/pdf_doc/fuzzy/fuzzy.pdf

Pathwork Diagnostics (2009d). Pathwork Diagnostics [On-line]. Available: http://www.pathworkdx.com/

Rosetta Genomics (2009e). Rosetta Genomics [On-line]. Available: https://www.rosettagenomics.com/index.asp

Asare, A. L., Ellis, J. S., & Caldwell, C. W. (2002). A decision-support system for flow cytometry immunophenotyping. *Am.J.Clin.Pathol., 118,* 567-573.

Bagg, A. (2005). Molecular diagnosis in lymphoma. *Curr.Hematol.Rep., 4,* 313-323.

Baird, K., Davis, S., Antonescu, C. R., Harper, U. L., Walker, R. L., Chen, Y. et al. (2005). Gene expression profiling of human sarcomas: insights into sarcoma biology. *Cancer Res., 65,* 9226-9235.

Bates, J. H. & Young, M. P. (2003). Applying fuzzy logic to medical decision making in the intensive care unit. *Am.J.Respir.Crit Care Med., 167,* 948-952.

Berner, E. & La Lande, T. (2007). Overview of Clinical Decision Support Systems. In E.Berner (Ed.), *Clinical Decision Support Systems: Theory and Practice* (Second ed., pp. 3-22). New York, NY: Springer.

Binaghi, E., Gallo, I., Ghiselli, C., Levrini, L., & Biondi, K. (2008). An integrated fuzzy logic and web-based framework for active protocol support. *Int.J.Med.Inform., 77,* 256-271.

Brown, R. E. (2009). Morphogenomics and Morphoproteomics. *Arch.Pathol.Lab Med., 133,* 568-579.

Cheang, M. C., van de Rijn, M., & Nielsen, T. O. (2008). Gene expression profiling of breast cancer. *Annu.Rev.Pathol., 3:67-97.,* 67-97.

Cheng, L., Zhang, S., MacLennan, G. T., Lopez-Beltran, A., & Montironi, R. (2009). Molecular and cytogenetic insights into the pathogenesis, classification, differential diagnosis, and prognosis of renal epithelial neoplasms. *Hum.Pathol., 40,* 10-29.

Cimino, J. J. (2007). Infobuttons and Point of Care Access to Knowledge. In R.Greenes (Ed.), *Clinical Decision Support: The Road Ahead* (First ed., pp. 345-371). Amsterdam: Elsevier.

Clarke, C. A., Glaser, S. L., Dorfman, R. F., Bracci, P. M., Eberle, E., & Holly, E. A. (2004). Expert review of non-Hodgkin's lymphomas in a population-based cancer registry: reliability of diagnosis and subtype classifications. *Cancer Epidemiol.Biomarkers Prev., 13,* 138-143.

Cook, E. H., Jr. & Scherer, S. W. (2008). Copy-number variations associated with neuropsychiatric conditions. *Nature., 455,* 919-923.

Costa, J. (2008). Is clinical systems pathology the future of pathology? *Arch.Pathol.Lab Med., 132,* 774-776.

Eddy, D. M. (2005). Evidence-based medicine: a unified approach. *Health Aff.(Millwood.), 24,* 9-17.

Elsheikh, T. M., Asa, S. L., Chan, J. K., DeLellis, R. A., Heffess, C. S., LiVolsi, V. A. et al. (2008). Interobserver and intraobserver variation among experts in the diagnosis of thyroid follicular lesions with borderline nuclear features of papillary carcinoma. *Am.J.Clin.Pathol., 130,* 736-744.

Farkas, D. H. (2008). Diagnostic molecular pathology in an era of genomics and translational bioinformatics. *Diagn.Mol.Pathol., 17,* 1-2.

Finn, W. G. (2007). Diagnostic pathology and laboratory medicine in the age of "omics": a paper from the 2006 William Beaumont Hospital Symposium on Molecular Pathology. *J.Mol.Diagn., 9,* 431-436.

Foran, D. J., Comaniciu, D., Meer, P., & Goodell, L. A. (2000). Computer-assisted discrimination among malignant lymphomas and leukemia using immunophenotyping, intelligent image repositories, and telemicroscopy. *IEEE Trans.Inf.Technol.Biomed., 4,* 265-273.

Foucar, E. (2001a). Classification in anatomic pathology. *Am.J.Clin.Pathol., 116 Suppl:S5-20.,* S5-20.

Foucar, E. (2001b). Diagnostic decision-making in anatomic pathology. *Am.J.Clin.Pathol., 116 Suppl:S21-33.,* S21-S33.

Friedman, B. A. (2008). A survey of the myriad forces changing anatomic pathology and their consequences. *Arch.Pathol.Lab Med., 132,* 735-738.

Geschwind, D. H. (2003). DNA microarrays: translation of the genome from laboratory to clinic. *Lancet Neurol., 2,* 275-282.

Glaser, J. & Hongsermeier, T. (2007). Managing the Investment in Clinical Decision Support. In R.Greenes (Ed.), *Clinical Decision Support: The Road Ahead* (First ed., pp. 403-422). Amsterdam: Elsevier.

Goldstein, L. J., Gray, R., Badve, S., Childs, B. H., Yoshizawa, C., Rowley, S. et al. (2008). Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. *J.Clin.Oncol., 26,* 4063-4071.

Greenes, R. (2007). Definition, Scope and Challenges. In R.Greenes (Ed.), *Clinical Decision Support: The Road Ahead* (First ed., pp. 3-29). Amsterdam: Elsevier.

Hess, J. L. (2002). The advent of targeted therapeutics and implications for pathologists. *Am.J.Clin.Pathol., 117,* 355-357.

Hicks, D. G., Kulkarni, S., & Hammond, M. E. (2008). The role of the indispensable surgical pathologist in treatment planning for breast cancer. *Arch.Pathol.Lab Med., 132,* 1226-1227.

Higgins, R. A., Blankenship, J. E., & Kinney, M. C. (2008). Application of immunohistochemistry in the diagnosis of non-Hodgkin and Hodgkin lymphoma. *Arch.Pathol.Lab Med., 132,* 441-461.

Hongsermeier, T., Kashyap, V., & Sordo, M. (2007). Knowledge Management Infrastructure: Evolution at Partners Healthcare System. In R.Greenes (Ed.), *Clinical Decision Support: The Road Ahead* (First ed., pp. 447-467). Amsterdam: Elsevier.

Hunt, J. L. (2009). Biomarkers in Anatomic Pathology: Adding Value to Diagnosis. *Arch.Pathol.Lab Med., 133,* 532-536.

Levenson, R. (2008). Putting the "more" back in morphology: spectral imaging and image analysis in the service of pathology. *Arch.Pathol.Lab Med., 132,* 748-757.

Marchionni, L., Wilson, R. F., Wolff, A. C., Marinopoulos, S., Parmigiani, G., Bass, E. B. et al. (2008). Systematic review: gene expression profiling assays in early-stage breast cancer. *Ann.Intern.Med., 148,* 358-369.

Marshall, C. R., Noor, A., Vincent, J. B., Lionel, A. C., Feuk, L., Skaug, J. et al. (2008). Structural variation of chromosomes in autism spectrum disorder. *Am.J.Hum.Genet., 82,* 477-488.

Maruya, S., Kim, H. W., Weber, R. S., Lee, J. J., Kies, M., Luna, M. A. et al. (2004). Gene expression screening of salivary gland neoplasms: molecular markers of potential histogenetic and clinical significance. *J.Mol.Diagn., 6,* 180-190.

Nakhleh, R. E. (2006). What is quality in surgical pathology? *J.Clin.Pathol., 59,* 669-672.

Nakhleh, R. E. (2008). Patient safety and error reduction in surgical pathology. *Arch.Pathol.Lab Med., 132,* 181-185.

Nguyen, A. N., De, J., Nguyen, J., Padula, A., & Qu, Z. (2008). A teaching database for diagnosis of hematologic neoplasms using immunophenotyping by flow cytometry. *Arch.Pathol.Lab Med., 132,* 829-837.

Papageorgiou, E. I., Spyridonos, P. P., Stylios, C. D., Ravazoula, P., Groumpos, P. P., & Nikiforidis, G. N. (2006). Advanced soft computing diagnosis method for tumour grading. *Artif.Intell.Med., 36,* 59-70.

Patel, V. L., Shortliffe, E. H., Stefanelli, M., Szolovits, P., Berthold, M. R., Bellazzi, R. et al. (2008a). The coming of age of artificial intelligence in medicine. *Artif.Intell.Med.*.

Patel, V. L., Shortliffe, E. H., Stefanelli, M., Szolovits, P., Berthold, M. R., Bellazzi, R. et al. (2008b). The coming of age of artificial intelligence in medicine. *Artif.Intell.Med.*.

Ramaswamy, S., Ross, K. N., Lander, E. S., & Golub, T. R. (2003). A molecular signature of metastasis in primary solid tumors. *Nat.Genet., 33,* 49-54.

Reis-Filho, J. S., Westbury, C., & Pierga, J. Y. (2006). The impact of expression profiling on prognostic and predictive testing in breast cancer. *J.Clin.Pathol., 59,* 225-231.

Rosai, J. (2001). The continuing role of morphology in the molecular age. *Mod.Pathol., 14,* 258-260.

Rosai, J. (2007). Why microscopy will remain a cornerstone of surgical pathology. *Lab Invest., 87,* 403-408.

Salto-Tellez, M. (2007). A case for integrated morphomolecular diagnostic pathologists. *Clin.Chem., 53,* 1188-1190.

Seising, R. (2006). From vagueness in medical thought to the foundations of fuzzy reasoning in medical diagnosis. *Artif.Intell.Med., 38,* 237-256.

Smith, M., Spence, M. A., & Flodman, P. (2009). Nuclear and mitochondrial genome defects in autisms. *Ann.N.Y.Acad.Sci., 1151:102-32.,* 102-132.

Sotiriou, C. & Pusztai, L. (2009). Gene-expression signatures in breast cancer. *N.Engl.J.Med., %19;360,* 790-800.

Weigelt, B., Horlings, H. M., Kreike, B., Hayes, M. M., Hauptmann, M., Wessels, L. F. et al. (2008). Refinement of breast cancer classification by molecular characterization of histological special types. *J.Pathol., 216,* 141-150.

Wick, M. R., Bourne, T. D., Patterson, J. W., & Mills, S. E. (2005). Evidence-based principles and practices in pathology: selected problem areas. *Semin.Diagn.Pathol., 22,* 116-125.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control, 8,* 338-352.

VITA

Gerald Arthur is a graduate of Carleton College and the University Of Chicago School Of Medicine. He served a rotating internship at Oregon Health Sciences University and a pathology residency at Dartmouth Medical School. He is certified by the American Board of Pathology in Anatomic and Clinical Pathology, Cytopathology and Hematopathology and practiced pathology for over 25 years as a member of Boyce and Bynum Pathology Laboratories in Columbia, Missouri. Upon retirement, he was very fortunate to obtain a National Library of Medicine Fellowship in biomedical informatics at the University of Missouri. Since completing the fellowship, he has been a research professor of pathology at the University Of Missouri School Of Medicine. Aside from decision support systems, Dr. Arthur's research interests include knowledge management, ontologies, the study of DNA regulatory regions and the epigenetics of lymphoma and leukemia.