Using ILIAD System Shell To Create an Expert System for Differential Diagnosis of Renal Masses

Phei Lang Chang, Yu-Chuan Li*, Chi Ju Wu, and Ming Hsiung Huang

Differential diagnosis of renal masses is an important and difficult process. A renal mass diagnostic system (RMDS) developed by using the ILIAD expert system shell has been created for diagnostic consultations and patient simulations. Seventy-two cases of renal mass have been tested on this system and the diagnostic accuracy was compared to that of residents. The overall diagnostic accuracy (75%) for renal masses is significantly better than second-year urological residents (60%) and not worse than urological chief residents (71%). The expert system also displays the cost of the diagnostic procedures so that the user can choose the most cost-effective diagnostic process. We conclude that this powerful renal mass diagnosis system developed by using ILIAD system shell can be used as a teaching, self-training and clinical tool for urological residents.

INTRODUCTION

An expert system is a type of artificial intelligence program used to solve special problems that normally require human expertise. Recently, a few successful medical expert systems have been developed to aid specialists solving domain-specific problems.[1–6]

Differential diagnosis of renal masses has been a difficult and expensive process for an urologist. However, this process is necessary before a renal mass operation because the position and characteristics of the mass may influence the surgical preparation, the surgical approach and the subsequent success of the operation. Therefore, we decided to use ILIAD expert system shell to create a differential diagnosis system that can help urologists in the preoperative work up of renal mass without performing redundant expensive or invasive procedures.

---

*From the Division of Urology, Department of Surgery, Chang Gung Memorial Hospital, Chang Gung Medical College, Taipei, Taiwan, Republic of China; and *Department of Medical Informatics, University of Utah, Salt Lake City, Utah.*

Table 1. An Example of Probabilistic Frame in RMDS

Disease: Renal Angiomyolipoma
Prevalence: 2 in 10,000 for inpatient surgical service population
Posterior Probability: 0.0

Findings	Cost	FPR	TPR	LR−	LR+
Abdominal CT scan shows renal mass with fat content	$401	0.001	0.95		1890
or					19.9
MRI of abdomen shows renal mass with high fat content	$663	0.001	0.95		1890
					19.9
Ultrasonography shows renal mass with high fat content	$143	0.005	0.80	170	
				6.63	
Abdominal CT scan shows renal mass with central bleeding	$401	0.005	0.95	949	19.9
or					
MRI of abdomen shows renal mass with central bleeding	$663	0.005	0.95	949	19.9
Ultrasonography shows renal mass with central bleeding	$143	0.01	0.85	85.0	
				6.60	
Renal angiography shows renal mass with neovascularity	$803	0.01	0.95	316	19.9
Intravenous pyelogram shows renal mass				7.00	
	$200	0.15	0.40	3.00	
Triad of tuberous sclerosis				5.00	
		0.01	0.30	1.34	
Abdominal examination: flank mass				3.00	
		0.15	0.30	1.28	
Present history: hypovolemic shock				3.00	
		0.10	0.45	1.28	
Abdominal examination: flank tenderness				2.50	
		0.15	0.20	1.20	
Evidence of anemia				1.33	
		0.15	0.20	1.06	

TPR (True Positive Rate) = Sensitivity
FPR (False Positive rate) = 1-Specificity
LR+: Positive likelihood ratio
LR−: Negative likelihood ratio

Table 1 are in U.S. dollars that come with ILIAD. This system also provides a function to change the cost of procedures as user's preference. We have created another version of the system using the actual costs in Taipei.

RMDS used estimated statistical associations between diseases and patient's clinical findings. The decision making information in RMDS was based on the knowledge and experience of urologists and scientific clinical literature in the field of urology. Prevalence rates (*a priori* probabilities) of the 18 renal mass diseases were calculated from the accumulated patient database of Chang Gung Memorial Hospital.

Each probability frame in the RMDS system includes *a priori* probability which represents the prevalence of the disease in the surgical service inpatient population and a posterior probability indicating the current probability of the disease given the patient

Table 2. An Example of Deterministic Frame in RMDS

Abnormal Urine Findings			
Findings	Status	Frequency	Cost
A. Urinalysis shows WBCs (>5/HPF)		0.05	$14
B. Urinalysis shows microscopic hematuria (RBC > 2/HPF in urine)		0.03	$14
C. Urinalysis shows bacteriuria		0.05	$14
True if A or B or C			

findings. The current status of each finding is indicated to the right of each finding (e.g., "Yes," "No," or "Unknown"). If the finding is a test, the relative cost will be shown underneath the status. The next column contains the findings true positive rate (sensitivity) and false positive rate (1-specificity). The final column shows the positive likelihood ratio and the negative likelihood ratio.

The differential diagnosis list and the probabilities associated with each disease can be displayed after the user presents the clinical findings of the patient. From the differential diagnosis list of RMDS, the user can select a disease and request to see the reasons that the specific diagnosis is being considered. This is a function inherently implemented in ILIAD system shell.[6,7]

Seventy-two consecutive cases of renal mass that we diagnosed and operated on between May 1989 and April 1992 were used to test the accuracy of the RMDS for discriminating the 18 renal masses that we implemented in this system. The final diagnoses of these 72 cases were all confirmed by pathological examination after operation and were used as the gold standard diagnoses in this study. The diagnoses made by the RMDS were determined as "correct" if the top diagnosis listed by the RMDS matched the gold standard diagnosis and the predicted probability of this diagnosis was greater than 50%. Four kinds of renal tumor including hemangiopericytoma, juxtaglomerular cell tumor, lipoma and cystadenocarcinoma could not be found in these test cases.

Three third year urological residents (the chief resident of urological department; numbered as CR-1, CR-2, and CR-3) and three second-year urological residents (numbered as R2-1, R2-2, and R2-3) evaluated these 72 cases. They were asked to make a single most likely diagnosis after reading the findings of those cases. They completed these test cases within 1 week. All six residents were informed that the results of the test would be known only to them and to the research team.

RESULTS

Of the 72 test cases, 31 had renal parenchymal tumor, 24 had renal pelvis tumor, 4 had renal cyst, 9 had renal abscess, and 4 had xanthogranulomatous pyelonephritis (XGP). Only one Wilms' tumor was found in these cases because children with Wilms' tumor were always admitted to the department of pediatric surgery. The number of patients with renal cyst did not correspond to its disease of prevalence because few renal cysts required exploration.

The RMDS was tested in these 72 cases undergoing surgical operation and produced a 75% overall diagnostic accuracy which compared with CR-1, CR-2, CR-3, R2-1, R2-2, and R2-3 who demonstrated 68%, 71%, 74%, 61%, 63%, and 57%. The mean diagnostic accuracy of 3 chief residents was 71% and the mean diagnostic accuracy of 3 second-year residents was 60%. The detailed results are shown in Table 3. In these 72 cases, 1 case of renal cell carcinoma and 2 cases of transitional cell carcinoma were diagnosed by the RMDS as the top diagnosis but with a posterior probability below 50%. Including these 3 cases, the overall diagnostic accuracy for renal masses of RMDS was 79%.

The diagnostic accuracy had no significant difference ($p > 0.05$) between RMDS and each chief resident, and no significant difference between RMDS and each R2 in diagnosis of renal parenchymal tumors or tumors of renal pelvis. However, the overall diagnostic accuracy for renal masses of RMDS was better than second-year urological residents with a significance level of $p < 0.05$ and not worse than chief residents. The 95% confidence intervals were determined by Pratt's approximation for binomial confidence limits (Table 4).[12,13] The diagnostic of RMDS and of residents are shown in Figure 1.

Table 3. The Number of Correct Diagnoses of Renal Masses made by RMDS and Urological Residents

	Cases	RMDS (%)	CR-1 (%)	CR-2 (%)	CR-3 (%)	R2-1 (%)	R2-2 (%)	R2-3 (%)
(1) Renal parenchymal tumors	31	21 (68)	23 (74)	23 (74)	24 (77)	22 (71)	16 (52)	16 (52)
Angiomyolipoma	4	3	4	3	3	2	2	3
Hemangiopericytoma	0							
Juxtaglomerular cell tumor	0							
Lipoma	0							
Lymphoblastoma	4	2	0	2	2	4	1	1
Metastatic tumor	2	1	2	2	2	2	1	1
Oncocytoma	4	1	0	0	1	0	0	0
Renal cell carcinoma	14	12	14	13	13	11	10	10
Sarcoma	2	1	2	2	2	2	1	0
Wilms' tumor	1	1	1	1	1	1	1	1
(2) Tumors of renal pelvis	24	18 (75)	16 (67)	16 (67)	17 (71)	15 (63)	16 (67)	14 (58)
Benign papilloma	3	3	3	3	2	3	2	1
Transitional cell carcinoma	17	14	13	13	14	12	14	13
Squamous cell carcinoma	2	1	0	0	1	0	0	0
Adenocarcinoma	2	0	0	0	0	0	0	0
(3) Renal cyst								
Simple cyst	4	4	2	3	3	2	2	3
Cystadenocarcinoma	0							
(4) Renal abscess	9	8	7	8	8	5	9	7
(5) XGP	4	3	1	1	1	0	2	1
Total	72	54 (75)	49 (68)	51 (71)	53 (74)	44 (61)	45 (63)	41 (57)

Table 4. The Statistical Analyses of Diagnostic Accuracy Among RMDS and Residents by Using Pratt's Approximation

N = 72	Correct cases	Percentage	CI* (Upper 95%)	CI* (Lower 95%)
RMDS[b]	54	75%	85%	63%
CR-1	49	68%	79%	56%
CR-2	51	71%	81%	59%
CR-3	53	74%	83%	62%
R2-1	44	61%	72%	49%
R2-2	45	63%	74%	50%
R2-3	41	57%	69%	45%

* Confidence intervals.
[b] The diagnostic accuracy of RMDS was better than second-year urological residents with a significance level of $p < 0.05$.

DISCUSSION

ILIAD is a frame based expert system for medical education that performs two major functions, consultations and simulations. In the consultation mode, an user presents a real case to ILIAD and ILIAD generates a differential diagnosis. In the simulation mode, ILIAD creates a simulated case and allows a user the opportunity to check his diagnostic skills.[14] The RMDS makes use of ILIAD system shell including functions of consultations and simulations to generate a differential diagnosis of renal mass and to provide residents with tools for doing systematic decision analysis and problem-based learning.

RMDS calculates each disease in its differential diagnosis list as an independent entity. The major difficulty of implementing this system is the estimation of the false positive rate of findings in each diagnosis. This value greatly affects the result of differ-

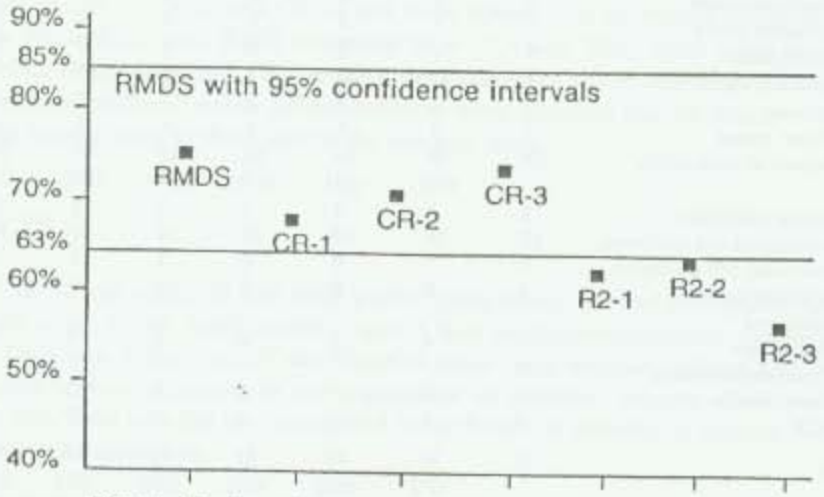

Figure 1. The diagnostic accuracy of RMDS and urological residents.

ential diagnosis. The false positive rate for findings is not regularly documented and is not a simple number for physicians to estimate. The expert judgment of an experienced urologist is often used to get an appropriate false positive rate.

We can use the RMDS to decide what would be the next most useful information for establishing a diagnosis of renal mass. By properly using the system, the diagnostic process can be optimized in terms of cost-effectiveness and physicians' preference. This function was based on the positive and negative likelihood ratio for each finding and the relative cost for tests.[9]

The differential diagnosis of renal mass was easily made in typical cases. For example, a diagnosis of renal abscess was suspected when positive percutaneous needle aspiration and a hypovascular renal mass in renal angiogram appeared in a patient with high fever. In atypical cases, differential diagnosis was very difficult. For example, renal oncocytoma without angiographic characteristic features was difficult to differentiate from renal cell carcinoma. In some rare cases, it was nearly impossible to make a preoperative differential diagnosis. For example, adenocarcinoma of renal pelvis cannot be differentiated preoperatively from transitional cell carcinoma without biopsy. Two of the test cases in this study were surgically confirmed as adenocarcinoma of renal pelvis, none of them were correctly diagnosed by either RMDS or the residents.

Some large tumors of the renal pelvis that invaded the parenchyma may mimic renal parenchymal tumors, revealing a big renal mass with a distorted caliceal system on the CT scan. It was very difficult to differentiate these two clinically, even though urine cytology was used. RMDS expert system tends to bring them up together in differential diagnosis list when the information put in supported one of them.

Certain renal tumors are sometimes accompanied by urinary tract infection, which make diagnosis doubtful. One situation that is commonly misdiagnosed as squamous cell carcinoma is the occurrence of xanthogranulomatous pyelonephritis combined with infected calculi, because they both present a tumor-like formation and renal function impairment.

RMDS was designed to make the user aware of the relative costs for various diagnostic procedures. Although some invasive diagnostic procedures carry valuable information for making diagnosis (e.g., renal angiography), this information may be available only at great cost. Therefore, in terms of cost effectiveness, these procedures may not be optimal at an early stage in the work-up. Other non-invasive diagnostic procedures may carry less diagnostic information (e.g., ultrasonography), but they can be less expensive and may be more appropriate in the early stages of the work-up. This function helps to remind user of performing non-invasive, low-cost diagnostic procedures at the early stage of working up a case.[9]

In this study, RMDS made correct diagnosis in all cases of benign papilloma of renal pelvis and renal simple cyst. In addition, RMDS identified three out of four cases of XGP while the residents averaged only one correct diagnosis. It appeared that RMDS might be most helpful for diagnosing diseases in certain categories, but on the other hand, we found no significant difference among RMDS, chief residents and the second-year residents while diagnosing specifically renal parenchymal tumors or tumors of renal pelvis. However, the overall diagnostic accuracy for renal masses of RMDS was significantly better than second-year residents and maintained the same accuracy level as chief residents.

In conclusion, a medical expert system can be easily created by using the ILIAD system shell. The combination of consultation and simulation functions can provide residents with an useful tool for self-training. RMDS also informed users about the costs of the diagnostic procedures so that the users can choose the most cost-effective way of working up the disease. The overall diagnostic accuracy for renal masses of RMDS was significantly better than second-year urological residents and not worse than chief residents. We believe that the overall diagnostic accuracy of RMDS will be better after revision of the system. However, in atypical cases of renal mass, RMDS tended to make more mistakes than the residents. A final check by an experienced urologist is still necessary for the presurgical diagnosis of renal masses.

ACKNOWLEDGMENTS

We are indebted to Homer R. Warner, MD, PhD, Peter J. Haug, MD, PhD, and other members of the Department of Medical Informatics and Applied Informatics Inc., University of Utah, for helpful suggestions concerning the manuscript and the ILIAD shell systems that made this project possible. We also thank Dr. John Kircher for the suggestions regarding statistical analysis.

REFERENCES

1. Donald, R., and Robert, H., Expert system in treating substance abuse. *West. J. Me.* 152:585–587, 1990.
2. Mira, E., Buizza, A., Magenes, G., Manfrin, M., and Schmid, R., Expert systems as a diagnostic aid in otoneurology. *ORL* 52:96–103, 1990.
3. Pryor, T., Gardner, R., Clayton, P., and Warner, H., The HELP System. *J Med. Syst.* 7:87–102, 1983.
4. Miller, R., INTERNIST-1/CADUCEUS. Problems facing expert consultant programs. *Methods Inf. Med.* 23:9–14, 1984.
5. Miller, R., Masarie, F., and Myers, J., Quick medical reference (QMR) for diagnostic assistance. *MD Comput.*3:34–48, 1986.
6. Warner, H., Haug, P., Lincoln, M., Warner, H., Jr., Sorenson, D., and Fan, C., Iliad as an expert consultant to teach differential diagnosis. In: Proceedings of the 11th Symposium on Computer Applications in Medical Care. Washington, DC: IEEE Computer Society Press, 1987, pp. 371–376.
7. Cundick, R., Turner, C., Lincoln, M., Buchanan, J., Anderson, C., Warner, H., Jr., and Bouhaddou, O., Iliad as a patient case simulator to teach medical problem solving. In: Proceedings of the 13th Symposium on Computer Applications in Medical Care, Washington, DC: IEEE Computer Society Press, 1989, pp. 902–906.
8. Turner, C., Lincoln, M., Haug, P., Warner, H., Williamson, J., and Whitman, N., Clustered disease findings: aspects of expert systems. International Symposium of Medical Informatics and Education. In (R. Salamon, D. Protti, and J. Moehr, eds.), University of Victoria, B.C., Canada, 1989, pp. 259–263.
9. Turner, C., Williamson, J., and Lincoln, M., et al. The effect of Iliad on medical student problem solving. In: Proceedings of the 14th symposium on Computer Applications in Medical Care, Washington, DC: IEEE Computer Society Press, 1990, pp. 478–481.
10. Lincoln, M., Turner, C., Hesse, B., and Miller, R., A comparison of clustered knowledge structures in Iliad and in Quick Medical Reference. In: Proceedings of the 12th symposium on Computer Applications in Medical Care, Washington, DC: IEEE Computer Society Press, 1988, pp. 131–135.
11. Bouhaddou, O., Warner, H., Yu, Hong, and Lincoln, M., The knowledge capabilities of the vocabulary

component of a medical expert system. In: Proceedings of the 14th symposium on Computer Applications in Medical Care, Washington, DC: IEEE Computer Society Press, 1990, pp. 655–660.
12. Bankowitz, R., McNeil, M., Challinor, S., Parker, R., Kapoor, W., and Miller, R., A computer-assisted medical diagnostic consultation service: Implementation and prospective evaluation of a prototype. *Ann. Intern. Med.* 110:824–832, 1989.
13. Blyth, C., Approximate binomial confidence limits. *J. Am. Stat. Assoc.* 81(395):843–855, 1986.
14. Yu, H., Haug, P., Lincoln, M., Turner, C., and Warner, H., Clustered knowledge representation: Increasing the reliability of computerized expert systems. In: Proceedings of the 12th symposium on Computer Applications in Medical Care, Washington, DC: IEEE Computer Society Press, 1988, pp. 126–130.