iRENEX: a clinically informed decision support system for the interpretation of 99mTc-MAG3 scans to detect renal obstruction

Ernest V. Garcia,
Department of Radiology and Imaging Sciences, Emory University School of Medicine, 1364 Clifton Rd, NE, Atlanta, GA 30322, USA Ernest.Garcia@emory.edu

Andrew Taylor,
Department of Radiology and Imaging Sciences, Emory University School of Medicine, 1364 Clifton Rd, NE, Atlanta, GA 30322, USA

Russell Folks,
Department of Radiology and Imaging Sciences, Emory University School of Medicine, 1364 Clifton Rd, NE, Atlanta, GA 30322, USA

Daya Manatunga,
Department of Radiology and Imaging Sciences, Emory University School of Medicine, 1364 Clifton Rd, NE, Atlanta, GA 30322, USA

Raghuveer Halkar,
Department of Radiology and Imaging Sciences, Emory University School of Medicine, 1364 Clifton Rd, NE, Atlanta, GA 30322, USA

Bital Savir-Baruch, and
Department of Radiology and Imaging Sciences, Emory University School of Medicine, 1364 Clifton Rd, NE, Atlanta, GA 30322, USA

Eva Dubovsky
Department of Radiology, University of Alabama at Birmingham, Birmingham, AL, USA

Abstract

Purpose—Decision support systems for imaging analysis and interpretation are rapidly being developed and will have an increasing impact on the practice of medicine. RENEX is a renal expert system to assist physicians evaluate suspected obstruction in patients undergoing mercaptoacetyltriglycine (MAG3) renography. RENEX uses quantitative parameters extracted from the dynamic renal scan data using QuantEM™II and heuristic rules in the form of a knowledge base gleaned from experts to determine if a kidney is obstructed; however, RENEX does not have access to and could not consider the clinical information available to diagnosticians interpreting these studies. We designed and implemented a methodology to incorporate clinical information into RENEX, implemented motion detection and evaluated this new comprehensive system (iRENEX) in a pilot group of 51 renal patients.

Methods—To reach a conclusion as to whether a kidney is obstructed, 56 new clinical rules were added to the previously reported 60 rules used to interpret quantitative MAG3 parameters. All the

---

clinical rules were implemented after iRENEX reached a conclusion on obstruction based on the quantitative MAG3 parameters, and the evidence of obstruction was then modified by the new clinical rules. iRENEX consisted of a library to translate parameter values to certainty factors, a knowledge base with 116 heuristic interpretation rules, a forward chaining inference engine to determine obstruction and a justification engine. A clinical database was developed containing patient histories and imaging report data obtained from the hospital information system associated with the pertinent MAG3 studies. The system was fine-tuned and tested using a pilot group of 51 patients (21 men, mean age 58.2±17.1 years, 100 kidneys) deemed by an expert panel to have 61 unobstructed and 39 obstructed kidneys.

Results—iRENEX, using only quantitative MAG3 data agreed with the expert panel in 87 % (34/39) of obstructed and 90 % (55/61) of unobstructed kidneys. iRENEX, using both quantitative and clinical data agreed with the expert panel in 95 % (37/39) of obstructed and 92 % (56/61) of unobstructed kidneys. The clinical information significantly ($p<0.001$) increased iRENEX certainty in detecting obstruction over using the quantitative data alone.

Conclusion—Our renal expert system for detecting renal obstruction has been substantially expanded to incorporate the clinical information available to physicians as well as advanced quality control features and was shown to interpret renal studies in a pilot group at a standardized expert level. These encouraging results warrant a prospective study in a large population of patients with and without renal obstruction to establish the diagnostic performance of iRENEX.

Keywords

Decision support systems; Renal obstruction; Clinical data; ^{99m}Tc-MAG3 renography

Introduction

Decision support systems (DSS) for imaging analysis and interpretation are rapidly being developed and will have an increasing impact on the practice of medicine. These DSS have the capacity to reduce costs, save physicians time, enhance quality control and improve diagnostic accuracy. These imaging DSS consist of computer software designed to assist physicians and other health-care professionals with image interpretation tasks, and include image processing, extraction of relevant variables, a dynamic (ever-expanding) medical knowledge base, an inference mechanism (engine) to reach a medical conclusion based on the knowledge base and patient data and a justification mechanism to justify its conclusions. Image analysis and interpretation, by itself, however, is incomplete since physicians typically have access to clinical information that may inform their scan interpretation.

Evaluation of suspected obstruction in adults using diuresis renography is an important clinical procedure [1–10]; however, the vast majority of studies are performed in institutions that perform fewer than three studies per week. To assist physicians with scan interpretation, particularly those with limited training or experience, we have previously developed [11] and validated [12] a renal expert system (RENEX) for detecting renal obstruction using ^{99m}Tc-mercaptoacetyltriglycine (MAG3) renal scans before and after administration of furosemide, at a standardized expert level. One important limitation of RENEX is that it is not capable of using the clinical information that might be available to physicians at the time of the study. In that trial, the output of RENEX was compared to the diagnosis of renal obstruction based on the interpretation of a panel of three experts who reviewed the scans for the presence or absence of renal obstruction without the benefit of any clinical information. To overcome this limitation, we initiated the development of iRENEX, an informed DSS that not only uses all the quantitative parameters extracted from the MAG3 renal study but also utilizes the same clinical information available to physicians to reach a conclusion regarding the presence or absence of obstruction.

Materials and methods

Patients

Renal studies from 51 patients (21 men, mean age 58.2±17.1 years; 100 kidneys. 51 right) were used as a pilot group to develop and test iRENEX. All studies used for this development were obtained from the renal database of patients referred to our nuclear medicine service to evaluate suspected renal obstruction. This study was performed under the purview and approval of Emory's Internal Review Board. Studies were selected from a subset of studies in our database that included a baseline ^{99m}Tc-MAG3 dynamic study followed by furosemide challenge in patients with a clinical history of other pertinent imaging/clinical procedures. This subset was selected from studies in patients more likely to have equivocal or abnormal results, since obstruction was excluded in patients with a normal baseline study, and patients with a normal baseline study did not undergo the second stage study with furosemide. In addition, studies were selected to include a variety of responses in order to develop a complete set of heuristic rules for interpreting renal obstruction.

Acquisition protocol

Patients were hydrated with 500 ml of water and positioned supine with the scintillation camera detector placed under the table. The study was conducted using a two-stage acquisition protocol which is a minor modification of the one-stage acquisition protocol recommended in the 1996 Santa Fe Consensus Report on diuresis renography [8]. A three-phase dynamic acquisition (baseline scan) was begun as approximately 370 MBq of ^{99m}Tc-MAG3 was injected. Phase one consisted of 24 2-s frames, phase two 16 15-s frames, and phase three 40 30-s frames. At the end of the baseline acquisition, one additional post-void 2-min image was obtained of the kidneys with the patient in the supine position, and 1-min anterior pre-void and post-void bladder images were also obtained to determine residual urine volume and post-void (30 min) over maximum (post-void/max) count ratios.

For all the patients in the study, review of the baseline scan could not exclude obstruction, and all patients in the study received an intravenous injection of furosemide. The majority of patients (86 %) received 40 mg of furosemide as recommended by recent consensus reports [8, 9] and furosemide administration was followed immediately by a second single-phase 20-min dynamic acquisition consisting of 40 30-s frames.

Data analysis

All patient studies were processed using QuantEM™ II , a new version of the QuantEM™ renal quantification program [13]. The QuantEM™ II software, developed specifically for ^{99m}Tc-MAG3, incorporates several quality control procedures to improve reproducibility [14], automatically detects patient motion [15], generates specific quantitative parameters recommended for scan interpretation and allows the MAG3 clearance to be calculated using a camera-based technique. The MAG3 clearance methodology contained in QuantEM™ has been previously extensively validated in a multicenter trial [16].

For the baseline renogram, a static image is summed from the frames from 2 to 3 min after injection. Using a filtered version of this image, whole kidney, background and cortical regions of interest (ROIs) are automatically defined. The user can override any of these automatic ROIs and replace them with manual ROIs. Background-subtracted curves are generated for the whole kidney and 47 quantitative parameters are generated including patient demographics (height, weight, age, sex, body surface area), curve parameters (time to peak counts, and 20 min to maximum count ratio for both whole kidney and cortical ROIs), voiding indices (post-void to pre-void and post-void to maximum count ratios) and the MAG3 clearance. These 47 parameters include all the parameters recommended in a

recent ISCORN consensus report for the performance and interpretation of diuresis renography studies [10]. The MAG3 clearance is calculated from the whole-kidney MAG3 counts from 1 to 2.5 min, and the preinjection and postinjection images of the dose syringe [16].

For the diuresis study, a static image is summed from the frames from 1 to 5 min after injection. ROIs are manually drawn for the whole kidney, background and renal collecting system. Background-subtracted curves are generated for the whole kidney and renal pelvis, and times-to-half-peak are calculated.

After processing the diuresis study, the baseline renogram results are loaded and QuantEM™ II calculates ratios comparing the first-minute counts and pre-void (last minute) counts in the diuretic study to the 1–2-min counts and peak counts in the baseline study. A database was created to hold the quantitative results of QuantEM™ II processing exported in XML format.

Patient motion and missing frame detection and reporting

A major quality control improvement added to iRENEX compared to the previous version RENEX is to automatically detect patient motion during acquisition of the baseline study and to use this information in modifying the certainty of obstruction. This algorithm has been reported elsewhere [15]. This is a two-step process. Briefly, QuantEM II detects the magnitude and time when the motion took place by using cross-correlation of count profiles of successive images in the dynamically acquired imaging sequence. The motion shift is calculated in shifts of 0.25 pixels (0.8 mm). iRENEX then converts these magnitude and time variables into certainty factors (CFs) which are used to modify the confidence of the determination of whether the kidney is obstructed.

Another improvement in quality control in iRENEX over RENEX is for QuantEM II to detect frames with missing counts near the end of the post-diuretic dynamic sequence and to pass this information to iRENEX which converts it to CFs and uses it in determining the certainty of obstruction. This count loss is usually due to premature termination of the study when the patient gets off the table to void.

Clinical database development

The clinical database was designed to contain data from patient histories and imaging reports obtained from the hospital information system (HIS). Records were entered in the clinical database over a period of weeks by manually transcribing data from the HIS. During this period the database design was revised and expanded in an iterative fashion. Each patient record consisted of 342 possible data entry fields.

Clinical variables used

The implemented clinical rules gleaned from human experts used one or more of the following variables present in the kidney at the time of the MAG3 study: (1) nephrostomy tube, (2) ureteral stent, (3) renal calculus, (4) ureteropelvic junction calculus, (5) obstructive calculus, (6) ureteral calculus, (7) hydronephrosis, (8) hydroureter, (9) ureteral stricture, (10) flank pain on arrival, (11) flank pain after diuretic administration, (12) renal artery stenosis, (13) pyeloplasty, and (14) abnormal creatinine level. In addition to patient history, the renal imaging reports considered were ultrasonography (last two), CT (last three), MRI, angiography, kidney-ureter-bladder (KUB) and retrograde contrast studies. The iRENEX algorithm included a temporal relationship test such that for each field the most abnormal finding was used if there was no recent information within 90 days of the MAG3 study, otherwise the most recent imaging data available were used ignoring fields that had no data.

Similarly when comparing clinical history information to this most recent imaging information, preference was given to the recent imaging information if it existed.

Expert panel review

Diagnosis of renal obstruction was based on the interpretation of a panel of three experts (A.T., E.D., R.H.) who reviewed the scans of all 51 patients in the pilot database for the presence or absence of renal obstruction. Each kidney was graded for the presence or absence of obstruction using the proposed MYCIN [17] CF scale, the same as that used by iRENEX. In this scale CFs range from –1 to +1, where –1 represents definitely not obstructed, –0.2 the cut-point between not obstructed and equivocal, +0.2 the cut-point between equivocal and obstructed, and +1 definitely obstructed. Each expert scored each kidney blinded to the results of the other experts and blinded to the results of iRENEX. After scoring each study for obstruction based on just the MAG3 quantitative renal parameters, each expert was unblinded to all the clinical information available for each patient for that study and rescored each study based on this additional information. The median score of the three scores from each of the experts was used as the final interpretation. When clinical information was considered together with the MAG3 study, the 100 kidneys from the 51 patient studies were deemed by the expert panel to have 61 unobstructed kidneys, 33 obstructed kidneys and 6 equivocal findings.

Decision support system

Previously 60 heuristic rules were extracted from the domain expert to generate the knowledge base for detecting obstruction from the renal quantitative data as described for RENEX [11]. The architecture of iRENEX is the same as that of RENEX with the addition of a quality control knowledge module that includes patient motion and missing frames and two additional clinical knowledge modules (for left and right kidneys) that take into account clinical information for an additional 56 newly incorporated rules.

Converting input parameters to CFs—The main difference between iRENEX and RENEX is the implementation of transformations for converting clinical findings to CFs. The reason for this difference is that most clinical findings (e.g. hydronephrosis) are not represented as continuous variables such as the T½, but rather as textual qualifiers. These clinical findings had to be transformed to CFs. If, for example, hydronephrosis was reported as present on a CT scan and described as mild, moderate or severe, the CF for the presence of hydronephrosis was assigned values of 0.2, 0.7 or 0.9, respectively. If hydronephrosis was present but not qualified by a descriptive adjective, the CF was 0.7. If the CT scan included the kidneys in the field of view, but there was no comment regarding hydronephrosis, the CF value was assigned as –0.2. Thus, these clinical variables were transformed to CFs as follows: not present –0.7, no data –0.2, equivocal 0.0, mild 0.2, moderate or present 0.7 and severe 0.9.

Knowledge base—The knowledge base was generated through systematic interviews to extract from the domain experts (A.T., E.D., R.H.) heuristic rules in an "IF A THEN B" format that are used by experts when they use specific renal parameters to reach a conclusion regarding whether or not a patient's kidney is obstructed. The domain experts were requested to provide both heuristic rules and the degree of certainty (CF) that the rule is believed to be true. All the clinical rules were implemented as metarules, meaning that they are considered after iRENEX reaches a preliminary conclusion on obstruction based on the quantitative MAG3 parameters, and the evidence of obstruction was then modified by the new clinical rules. For example, a typical clinical rule reads: IF there is evidence this kidney has a hydroureter and evidence that this kidney's MAG3 clearance is abnormal THEN there is a moderate (CF= 0.3) evidence that this kidney is obstructed.

Statistical analysis

The median CF of the three CFs from each of the experts with their use of clinical information was used as the final interpretation, and thus the reference standard. All studies were classified as obstructed (CF 0.2), unobstructed (CF –0.2), or equivocal (–0.2< CF <0.2). To categorize the agreements between the experts' consensus and the iRENEX conclusions, analysis was performed with the equivocal CFs deemed not obstructed and repeated with the equivocal CFs deemed obstructed for both the experts' consensus and iRENEX conclusions.

The difference between two population proportions from a single sample [18] was used to test if there were differences in interpretation agreement between iRENEX with the benefit of the new clinical rules and iRENEX without the benefit of the new clinical rules compared to the reference standard.

Categorical variables are expressed in simple proportions. Spearman's correlation coefficient was used to compare the experts' CFs with the iRENEX CFs for interpreting obstruction with the benefit of clinical information. Similarly, the certainty of iRENEX in detecting obstruction using just the QuantEM™ II quantitative parameters was compared to the certainty of using these parameters plus the pertinent clinical information provided. Statistical significance of these comparisons were determined using two-tailed paired t-tests. A $p<0.05$ was considered significant for all comparisons.

Results

The results associated with this development include: (1) the boundary conditions used to transform the new clinical parameter values to CFs to form the updated parameter knowledge library, (2) the heuristic rules gleaned from the domain experts and modified by the pilot group which formed the updated knowledge base, and, once the parameter knowledge library and knowledge base were updated, were optimized to the pilot group, (3) analysis of the agreement between iRENEX and the human experts in the pilot group as to whether kidneys were obstructed, (4) analysis of the agreement of the degree of certainty of kidney obstruction between the experts and iRENEX expressed as the correlation between the consensus of the experts and the CF from iRENEX, and (5) analysis of the certainty by iRENEX in detecting obstruction using the MAG3 quantitative variables alone compared to using these quantitative variables plus the clinical information provided.

Table 1 shows the effects of motion and missing frame rules on the pertinent quantitative variables. So for example, the first entry in the table signifies that detection of motion on the baseline MAG3 study (before furosemide administration) would signal that the CF associated with "whole kidney half-time is abnormal" is artificially decreased since the motion makes the kidneys misregister the ROI and gives the appearance that the kidney clears more quickly than it actually occurs. To counteract this artificial decrease in CF, the certainty of the whole-kidney half-time is increased (shifted toward a more abnormal value). Table 2 shows how a specific rule affects the likelihood (CF) of obstruction. For example, the presence of hydronephrosis slightly increases the likelihood of obstruction (CF 0.1). Both tables show the magnitude and direction (positive or negative) of the evidence inferred and the frequency that each rule is used in the evaluation of the 100 kidneys in the pilot group of 51 patients. To use a rule means that the IF part of the rule was found to be true so the THEN part of the rule was executed. Not shown are the results from a rule related to the furosemide dose which reads: If the kidney function is abnormal (or creatinine level is elevated) and the kidney is obstructed and the diuretic dose is 50 mg then there is moderate negative evidence (CF –0.2) of obstruction. This rule was used 28 times in interpreting the 100 kidneys in the pilot group.

Table 3 shows the agreement between iRENEX and the median reading of the three human experts used as consensus as to whether or not the kidneys were obstructed. The iRENEX analysis was repeated with and without the benefit of the new clinical rules. The analysis was also performed using equivocal interpretations as negative for obstruction and repeated as positive for obstruction. When equivocal was deemed positive for obstruction, the DSS iRENEX, using only quantitative MAG3 data, agreed with the expert panel in 87 % (34/39) of the obstructed kidneys and in 90 % (55/61) of the unobstructed kidneys. Using both quantitative MAG3 and clinical data, iRENEX agreed with the expert panel in 95 % (37/39) of the obstructed kidneys and in 92 % (56/61) of the unobstructed kidneys. The improvement in the agreement on the presence of obstruction for iRENEX with the benefit of the clinical information and rules compared to iRENEX without the clinical information was statistically significant (p=0.04).

Figure 1 shows the correlation between iRENEX's CFs and the median CF of the three human experts used as consensus as to whether or not the kidneys were obstructed. Note the excellent correlations for both the left (R=0.88, y=0.90x+0.02) and right (R=0.86, y.93x–0.01) kidneys, confirming that the confidence in the iRENEX conclusions matched the confidence in the human experts' consensus conclusions. The iRENEX scores and the experts' scores were not significantly different (p=0.07 for the left kidney, p=0.85 for the right kidney; paired t-test). The processing time per patient was practically instantaneous using a 3.0 GHz PC programmed using IDL.

Finally, the iRENEX results agreed with the experts' interpretation of obstruction with significantly higher certainty (CF=0.52±0.29) than the RENEX results using the quantitative variables alone (CF=0.43±0.35; p<0.001). Since the clinical rules were mostly not used for detecting the absence of obstruction, the difference in the certainty of the interpretation of the absence of obstruction with and without the addition of the clinical rules (CF=–0.70±0.21 and CF = –0.70±.0.22, respectively) was not significant (p=0.45).

Discussion

Our previously developed RENEX system only used quantitative parameters extracted from MAG3 renal studies to determine whether or not a kidney was obstructed [11, 19]. In this project we have substantially expanded RENEX to use the same clinical information available to the clinician to supplement the scan interpretation. Although there are 342 fields for each patient, the clinician only enters the data available for the interpretation of the study; the remaining fields are set to default values. In the future, data mining and natural language processing should populate these fields automatically. Our approach consisted of developing: (1) a clinical database to contain data from patient renal histories and imaging reports obtained from the HIS and to relate these reports to the pertinent MAG3 studies, (2) an updated parameter knowledge library with the list of the boundary conditions necessary for transforming the values of each clinical parameter to a CF, and (3) an updated knowledge base of 116 heuristic rules used to reach conclusions regarding the image interpretation.

There was excellent agreement between iRENEX and the consensus reading of three experts as to whether or not there was obstruction of the kidneys of the pilot group that was also used as a training set to develop iRENEX. Agreement would likely have been even higher if all subjects referred for suspected obstruction had been included in the pilot group. In practice, obstruction is excluded in about one-third of patients by a normal baseline scan, and this group of obviously non-obstructed patients did not receive furosemide and was purposely not included in the pilot training group. There was a significant improvement in detecting obstruction with the use of clinical information by iRENEX compared to using the

quantitative scan parameters alone, although the excellent agreement between iRENEX and the experts was consistent whether or not clinical information routinely available to physicians was used. There was also excellent agreement between the experts and iRENEX in the degree of certainty as to whether the kidney was obstructed.

It is important to note that these results applied only to patients in the training set, and do not validate the method. The training set was utilized to optimize iRENEX, and a prospective study is needed for validation. Nevertheless, our results incorporating clinical information are encouraging. We utilized a similar approach in the analysis of the quantitative data obtained from the scans by first evaluating a pilot group and then confirming those results in a prospective group of 95 patients (185 kidneys) where, using quantitative data only, RENEX agreed with the consensus expert reading in 84 % (101/120) of unobstructed kidneys, in 92 % (33/36) of obstructed kidneys, and in 45 % (13/29) of equivocal kidneys [11, 12], We subsequently demonstrated that the interpretation of RENEX was actually equivalent to that of an expert [12]. The main difference between that prospective trial [12, 20] and our present study was that clinical data was not considered by either the experts or the expert system. The present iRENEX results are encouraging, and a large prospective trial to validate the approach is warranted. To our knowledge there are no computer-aided diagnostic tools to aid clinicians in the interpretation of renal scans to determine renal obstruction.

We did not use clinical outcome as the gold standard. Clinical outcome is an attractive goal but it misses the point of an expert system which is to interpret studies with the same level of expertise as experts. It is generally accepted that experts interpret studies in their specialty better than general radiologists. This is the basis for having distinct areas of expertise within academic departments and private practice settings. Outcome is certainly an important measure, but in diuresis renography outcome as a gold standard is confounded by the fact the scan interpretation (obstruction versus no obstruction) has a major impact on outcome (surgical intervention versus observation); consequently, this gold standard is biased. Other factors further confound outcome as a gold standard; for example, a patient having a pyeloplasty to relieve obstruction 1 year after a diuresis renography scan was interpreted as "no obstruction". In this example, did the scan miss obstruction a year earlier, did the scan indicate obstruction but did the imager overlook the critical findings and interpret the scan incorrectly, did the patient only become obstructed 1 year following the scan or did an aggressive surgeon operate on a non-obstructed kidney? Using patient outcome as a gold standard has an inherent bias, interpretation of the results is not straightforward and it is not the goal of an expert system.

Based on a traditional understanding of the role of clinical information, the addition of clinical information might have been expected to have a greater impact on the final interpretation than we observed, particularly in view of the number of clinical rules used by iRENEX in reaching an interpretation as shown in Table 2. Several factors may have contributed to these results. First, agreement with the experts without clinical information was already quite good; consequently, it would have been difficult to show a large improvement by the addition of clinical information. Second, iRENEX was designed such that a conclusion is reached based on the scan parameters alone; this conclusion is then modified by the use of clinical rules. Although a large number of clinical rules were used by the iRENEX interpretation (Table 2), the effect of the particular rules that were used was modest and not sufficient to shift the interpretation from one diagnostic category to another. Nevertheless, we did show that the clinical information provided to iRENEX significantly ($p<0.001$) increased its certainty in correctly detecting obstruction over using the quantitative data alone.

We are continuing to investigate how to further improve the diagnostic performance of iRENEX. We have noted from the present pilot study a discrepancy between iRENEX and the human experts when there is a calyceal obstruction in a localized kidney region. Resolution of this difference will require the development of image analysis routines that determine the regional status of each kidney. Another important limitation of this version of iRENEX is that its entire knowledge base refers to adult renal parameters. A separate knowledge base is required for implementing this approach in children which would have to take into account the differences in existing knowledge, and will also require a totally different patient population [21–23]. Finally, the three-phase acquisition protocol used to acquire data for the archived patients randomly selected for this study was initiated in the early 1990s and was a compromise based on limitations of memory and processing speed. The three-phase acquisition interval is not optimal for Patlak-Rutland plots [24] or cortical transit times. Parameters derived from these approaches may augment the discrimination between obstructed and non-obstructed kidneys, and we plan to test this possibility in the future using our current acquisition protocol of 2 s per frame for 5 min followed by 10 s per frame for 19 min.

Conclusion

We have substantially expanded a renal expert system previously developed for detecting renal obstruction using ^{99m}Tc-MAG3 renal scans before and after furosemide administration by incorporating quality control features, motion detection and the clinical information that would be available to clinicians at the time of scan interpretation. In addition to developing a methodology for incorporating clinical information in imaging expert systems, we have shown that, in a preliminary pilot group of patients, iRENEX interprets these studies at a high level of agreement with experts and with significantly increased certainty when using the available clinical data. On the basis of these encouraging preliminary results, a prospective study in a large population of patients with and without renal obstruction is warranted to establish the diagnostic performance of this expanded renal expert system.

Acknowledgments

References

1. Woolfson RG, Neild GH. The true clinical significance of renography in nephro-urology. Eur J Nucl Med. 1997; 24:557–70. [PubMed: 9142738]

2. Kletter K, Nurnberger N. Diagnostic potential of diuresis renography: limitations by the severity of hydronephrosis and by impairment of renal function. Nucl Med Commun. 1989; 10:51–61. [PubMed: 2919032]

3. Upsdell SM, Leeson SM, Brooman PJ, O'Reilly PH. Diuretic-induced urinary flow rates at varying clearances and their relevance to the performance and interpretation of diuresis renography. Br J Urol. 1988; 61:14–8. [PubMed: 3342295]

4. Kuyvenhoven J, Piepsz M, Ham H. When could the administration of furosemide be avoided? Clin Nucl Med. 2003; 28:732–7. [PubMed: 12972994]

5. Brown SC, Upsdell SM, O'Reilly PH. The importance of renal function for the interpretation of diuresis renography. Br J Urol. 1992; 69:121–5. [PubMed: 1537020]

6. Chaiwatanarat T, Padhy AK, Bomanji JB, Nimmon CC, Sonmezoglu K, Britton KE. Validation of renal output efficiency as an objective quantitative parameter in the evaluation of upper urinary tract obstruction. J Nucl Med. 1993; 34:845–8. [PubMed: 8478722]

7. Schlotmann A, Clorius JH, Clorius SN. Diuretic renography in hydronephrosis: renal tissue tracer transit predicts functional course and thereby need for surgery. Eur J Nucl Med Mol Imaging. 2009; 36:1665–73. doi:10.1007/s00259-009-1138-5. [PubMed: 19437014]

8. O'Reilly P, Aurell M, Britton K, Kletter K, Rosenthal L, Testa T. Consensus on diuresis renography for investigating the dilated upper urinary tract. J Nucl Med. 1996; 37:1872–6. [PubMed: 8917195]

9. O'Reilly PH, Consensus Committee of the Society of Radionuclides in Nephrology. Standardization of the renogram technique for investigating the dilated upper urinary tract and assessing the results of surgery. BJU Int. 2003; 91:239–43. [PubMed: 12581012]

10. Taylor AT, Blaufox MD, De Palma D, Dubovsky EV, Erba B, Eskild-Jensen A, et al. Guidance document for structured reporting of diuresis renography. Semin Nucl Med. 2012; 42(1):41–8. [PubMed: 22117812]

11. Garcia EV, Taylor A, Halkar R, Folks R, Krishnan M, Cooke CD, et al. RENEX: an expert system for the interpretation of 99mTc-MAG3 scans to detect renal obstruction. J Nucl Med. 2006; 47(2): 320–9. [PubMed: 16455639]

12. Taylor A, Garcia EV, Binongo JN, Manatunga A, Halkar R, Folks RD, et al. Diagnostic performance of an expert system for interpretation of 99mTc MAG3 scans in suspected renal obstruction. J Nucl Med. 2008; 49:216–24. [PubMed: 18199609]

13. Taylor A Jr, Corrigan PL, Galt J, Garcia EV, Folks R, Jones M, et al. Measuring technetium-99m-MAG3 clearance with an improved camera-based method. J Nucl Med. 1995; 36:1689–95. [PubMed: 7658232]

14. Folks RD, Garcia EV, Taylor AT. Development and prospective evaluation of an automated software system for quality control of quantitative 99mTc-MAG3 renal studies. J Nucl Med Technol. 2007; 35:27–33. [PubMed: 17337654]

15. Folks RD, Manatunga D, Garcia EV, Taylor AT. Automated patient motion detection and correction in dynamic renal scintigraphy. J Nucl Med Technol. 2011; 39:131–9. [PubMed: 21565959]

16. Taylor A, Manatunga A, Morton K, Reese L, Prato FS, Greenberg E, et al. Multicenter trial validation of a camera-based method to measure Tc-99m mercaptoacetyltriglycine, or Tc-99m MAG3, clearance. Radiology. 1997; 204:47–54. [PubMed: 9205222]

17. Shortliffe, EH. Computer-Based Medical Consultations: MYCIN. Elsevier Scientific; Amsterdam: 1976. p. 264

18. Dunn, OJ. Basic statistics: a primer for the biomedical sciences. Wiley; New York: 1977. p. 116-9.

19. Garcia EV, Taylor A, Manatunga D, Folks R. A software engine to justify the conclusions of an expert system for detecting renal obstruction on 99mTc-MAG3 scans. J Nucl Med. 2007; 48:463–70. [PubMed: 17332625]

20. Binongo JN, Manatunga A, Taylor AT. Computer-aided diagnosis of renal obstruction: utility of log-linear modeling versus standard ROC and kappa analysis. EJNMMI Res. 2011; 1:1–8. [PubMed: 21935501]

21. Gordon I, Colarinha P, Fettich J, Fischer S, Frökier J, Hahn K, et al. Guidelines for standard and diuretic renography in children. Eur J Nucl Med. 2001; 28:BP21–30. [PubMed: 11315613]

22. Eskild-Jensen A, Gordon I, Piepsz A, Frøkiær J. Interpretation of the renogram: problems and pitfalls in hydronephrosis in children. BJU Int. 2004; 94:887–92. [PubMed: 15476528]

23. Howman-Giles R, Uren R, Roy LP, Filmer RB. Volume expansion diuretic renal scan in urinary tract obstruction. J Nucl Med. 1987; 28:824–8. [PubMed: 3553462]

24. Peters AM. Graphical analysis of dynamic data: The Patlak-Rutland plot. Nucl Med Commun. 1994; 15:669–72. [PubMed: 7816377]
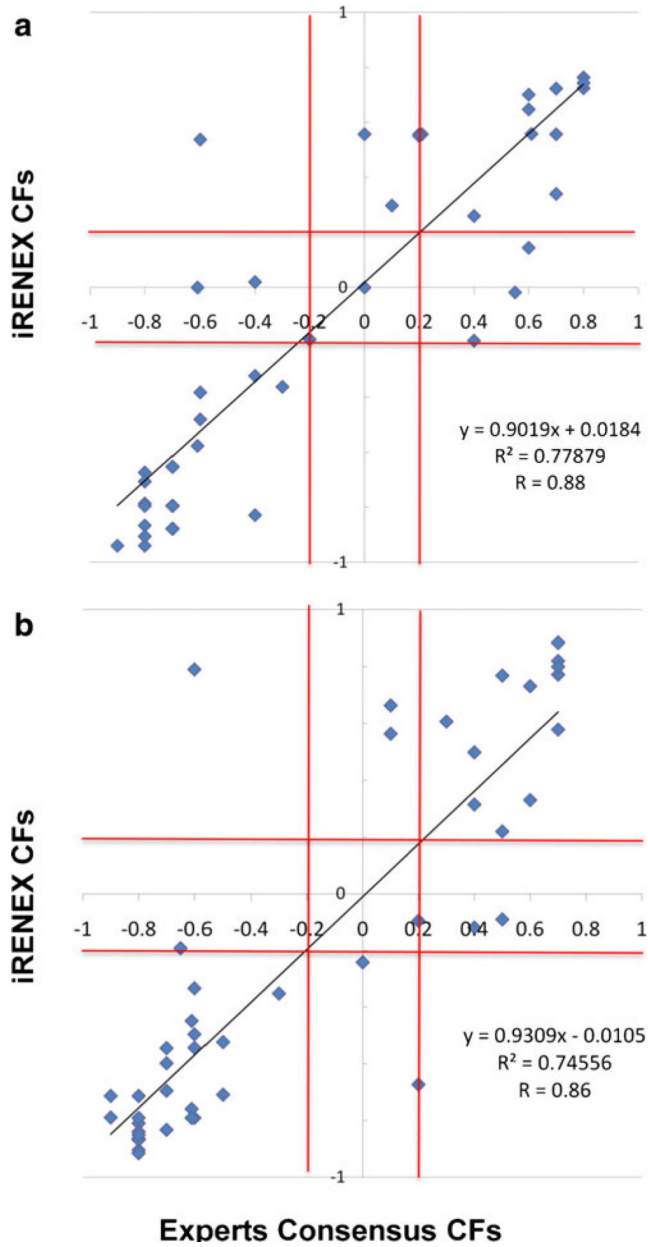
Fig. 1.
Correlations between iRENEX CFs and the median CF of the three human experts used as consensus as to whether or not the kidneys were obstructed: **a** left kidney; **b** right kidney. *Red lines* equivocal boundaries CF=–0.2 and CF=0.2. Note the excellent degree of correlation, confirming that iRENEX conclusions show a very similar degree of certainty to the experts' conclusions

Table 1

Effect of motion rules on quantitative variables

Rule	Variable affected	Evidence inferred	Frequency of use[b]
Baseline late motion[a]	Whole-kidney T½	Very strong positive evidence[b]	12
Baseline late motion[a]	20 min to max ratio	Very strong positive evidence[b]	12
Furosemide missing frames	Pre-void to baseline maximum	Equivocal	4

[a] Late motion inappropriately reduces the counts in the kidney ROI. Consequently, the CFs associated with these variables will appear artificially more normal than they should be. Applying the positive evidence of the motion rules will help correct this error and move the CFs towards the equivocal range.

[b] Frequency of use means how many times a rule was used in processing all the patients. Using a rule means that the IF part of the rule was found to be true so the THEN part of the rule was executed which modified the certainty of the interpretation

Table 2

Effect of clinical rules on iRENEX detecting obstruction

Rule	Inferred evidence of obstruction	CF	Frequency of use[a]
Hydronephrosis	Very mild positive evidence	0.1	39
Hydronephrosis and abnormal function	Very mild positive evidence	0.1	27
Hydronephrosis and hydroureter	Very mild positive evidence	0.1	12
Ureteral stent	Mild negative evidence	–0.2	10
Hydroureter and abnormal function	Very mild positive evidence	0.1	8
Renal calculus	Mild positive evidence	0.2	6
Ureteral calculus	Mild positive evidence	0.2	5
Flank pain on arrival	Very mild positive evidence	0.1	5
Ureteral stricture	Moderate positive evidence	0.3	2
Nephrostomy	Mild negative evidence	–0.2	2
Pyeloplasty	Moderate negative evidence	–0.3	2
Negative hydronephrosis	Mild negative evidence	–0.2	2

[a]Frequency of use means how many times a rule was used in processing all the patients. Using a rule means that the IF part of the rule was found to be true so the THEN part of the rule was executed which modified the certainty of the interpretation. For example, the Hydronephrosis row means that of the 100 kidneys, 39 where found to have hydronephrosis from the clinical information resulting in increasing the evidence of obstruction in each kidney by a certainty of 0.1 (very mild).

Table 3

iRENEX agreement with human experts

Criteria[a]	Equivocal scores	Left kidney			Right kidney		
		True-positive	True-negative	Agreement[b]	True-positive	True-negative	Agreement[b]
Quantitative	Considered not obstructed	11/14 (79%)	29/35 (80%)	40/49 (82%)	9/14 (64%)	34/37 (92%)	43/51 (84%)
	Considered obstructed	20/20 (100%)	25/29 (86%)	45/49 (94%)	14/19 (74%)	30/32 (94%)	44/51 (86%)
Quantitative + clinical	Considered not obstructed	13/14 (92 %)	29/35 (80%)	42/49 (86%)	12/14 (86%)	34/37 (92%)	46/51 (90%)
	Considered obstructed	20/20 (100%)	26/29 (90%)	46/49 (94%)	17/19 (89%)	30/32 (94%)	47/51 (92%)

The results were not significantly different between the left and right kidneys ($p>0.05$).

[a] *Quantitative* quantitative MAG3 variables only; *Clinical* clinical and quality control variables.

[b] Agreement between iRENEX and expert.