Published in final edited form as:

AJR Am J Roentgenol. 2007 May ; 188(5): 1395–1402. doi:10.2214/AJR.06.0931.

Evaluation of Two Diuresis Renography Decision Support Systems to Determine the Need for Furosemide in Patients with Suspected Obstruction

Andrew Taylor¹, Andrew N. Hill², José N. E. Binongo², Amita K. Manatunga², Raghuv eer Halkar¹, Eva V. Dubovsky³, and Ernest V. Garcia¹

¹Department of Radiology, Division of Nuclear Medicine, Emory University School of Medicine, 1364 Clifton St., Atlanta, GA 30322

²Department of Biostatistics, Emory University School of Public Health, Atlanta, GA

³Department of Radiology, Division of Nuclear Medicine, University of Alabama School of Medicine, Birmingham, AL

Abstract

OBJECTIVE—The purpose of this study was to compare the decisions regarding the need for furosemide made by two independent renal decision support systems, RENEX and CARTAN, with the need for furosemide determined in clinical practice and by expert reviewers using the baseline plus furosemide protocol.

SUBJECTS AND METHODS—RENEX and CARTAN are independent decision support systems that reach their conclusions without operator input. RENEX is a knowledge-based system and CARTAN is a statistical decision support system. Both were trained using the same pilot group of 31 adult patients (61 kidneys) referred for suspected obstruction. Subsequently, both systems were prospectively applied to 102 patients (200 kidneys) of whom 70 received furosemide; decisions regarding the need for furosemide were compared with the clinical decisions and the decisions of three experts who independently scored each kidney on the need for furosemide. Differences were resolved by consensus.

RESULTS—RENEX agreed with the clinical and experts' decisions to give furosemide in 97% (68/70) and 98% (65/66) of patients, respectively, whereas CARTAN agreed in 90% (63/70) and 89% (59/66), respectively, $p < 0.03$. In contrast, CARTAN agreed with the experts' decision to withhold furosemide in 78% of kidneys (87/111), whereas RENEX agreed in only 69% of kidneys (77/111), $p = 0.008$.

CONCLUSION—Use of RENEX or CARTAN as decision support tools in the baseline plus furosemide protocol has the potential to help the radiologist avoid unnecessary imaging and reduce the technologist, computer, camera, and physician time required to perform the procedure.

Keywords

genitourinary tract imaging; kidney; nuclear medicine; radiology practice; renal disease

In the evaluation of suspected ureteral obstruction, an international consensus panel recommended baseline imaging with technetium-99m mercaptoacetyltriglycine (^{99m}Tc

MAG3) followed by furosemide administration and 15 minutes of additional imaging [1]. Some investigators have recommended strategies to eliminate the need for furosemide in selected patients [2]. The authors of the consensus report did note that baseline imaging might be sufficient to exclude obstruction in experienced hands, but this option was not their general recommendation [1]. Nevertheless, if the baseline examination can exclude obstruction, the patient can be spared the administration of furosemide and additional imaging time, and there will also be a cost saving because of the reduction of camera, computer, and technologist time required to complete the furosemide component of the study and physician time required to interpret it.

Despite these advantages, it may be challenging for a general radiologist to acquire the experience and expertise in diuresis renography to determine with confidence whether a kidney is obstructed or when furosemide is not required. Radiologists are required to master an ever-expanding knowledge base while the hours available to master this knowledge base and apply it to specific tasks are steadily shrinking because of pressure to increase the number of studies each radiologist interprets. The convergence of an expanding knowledge base and escalating time constraints inevitably increases the risk of physician error. This is particularly true of low-volume studies such as diuresis renography in which radiologists may have had limited training and experience. In fact, a large percentage of the estimated 590,000 renal imaging examinations performed annually in the United States are interpreted at sites that perform fewer than three studies per week [3]. For these reasons, it is particularly important to develop and implement decision support tools to assist physicians in interpreting low-volume studies so that such studies can be interpreted at a faster rate and at a higher level of expertise.

The digital and dynamic nature of ^{99m}Tc MAG3 renal scanning data has made it possible to quantify a variety of functional renal parameters to produce an expanded knowledge base. Two independent renal decision support systems (RENEX and CARTAN) are being developed to use this expanded knowledge base to determine the need for furosemide and to detect renal obstruction in patients referred for diuresis renography [4, 5]. RENEX (renal expert) uses a set of rules obtained from human experts to analyze the quantitative parameters obtained from MAG3 scintigraphy, whereas CARTAN (classification and regression tree analysis) is a statistical method that grows and prunes a decision tree on the basis of an analysis of these quantitative parameters in a training data set [6].

As an intermediate step to developing a decision support system to detect obstruction, this prospective study was performed to compare the decisions regarding the need for furosemide made by a heuristic approach (RENEX) and an analytic approach (CARTAN) with the need for furosemide determined in clinical practice and by expert reviewers.

Subjects and Methods

This study was performed under the purview and with the approval of the internal review board at Emory University School of Medicine. Data collection and database use were compliant with the terms of the Health Insurance Portability and Accountability Act (HIPPA).

Acquisition Protocol and Data Collection

Patients were hydrated with approximately 10 oz (296 mL) of water on arrival in the department. Imaging was performed with the patient supine and the scintillation camera detector placed under the table. A three-phase dynamic acquisition was begun at the time of injection of approximately 10 mCi (370 MBq) of ^{99m}Tc MAG3. Phase one consisted of 24 2-second frames, phase two was 16 15-second frames, and phase three was 40 30-second

frames. All patient studies were processed using the QuantEM 2.0 renal quantification program (developed at Emory University, licensed to GE Healthcare) to generate the input parameters for RENEX and CARTAN.

The QuantEM software, developed specifically for ^{99m}Tc MAG3 [7], has been validated in a multi-center trial [8], incorporates several quality control procedures to improve reproducibility, and generates specific quantitative parameters recommended for scan interpretation.

To process the baseline renogram, a static image is summed from the 2- to 3-minute postinjection frames. 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 whole-kidney and cortical curves are generated, and 47 quantitative parameters are generated including patient demographics (height, weight, age, sex, body surface area), curve parameters (time-to-peak counts and 20-minute-to-maximumcount ratio for both whole-kidney and cortical ROIs), voiding indexes (postvoid to prevoid and postvoid to maximum count ratios), relative uptake, and the MAG3 clearance. The MAG3 clearance is calculated from the 1- to 2.5-minute whole-kidney uptake of MAG3 corrected for renal depth and attenuation and the preinjection and postinjection images of the dose syringe [7–10].

Protocol Survey

To estimate the frequency of use of the baseline plus furosemide protocol, one of the authors asked for a show of hands or keypad response from radiologists and nuclear medicine physicians attending renal lectures at the Educational Symposia Nuclear Medicine course in March 2005, the Emory course in Nuclear Medicine in July 2005, the 2005 Northeast Regional Meeting of the Society of Nuclear Medicine (SNM), and a 2005 educational course at the Radiological Society of North America (RSNA).

RENEX

The architecture of RENEX was inspired by two previously developed expert systems, MYCIN (Stanford University) [11] and PERFEX (Georgia Institute of Technology) [12], and is the subject of a separate publication [5]. MYCIN is a pioneering rule-based expert system developed in the 1970s to help physicians determine the appropriate antibiotic for patients with infections. The name MYCIN was chosen because many of the available antibiotics included “mycin” in the name of the antibiotic. PERFEX is a commercially available imaging expert system developed to assist physicians in the interpretation of myocardial perfusion SPECT studies [13].

Normal limits were established for 47 quantitative parameters extracted from the ^{99m}Tc MAG3 scans of 106 potential renal donors [14, 15]. These 47 parameters consisted of seven demographic patient parameters, 20 left kidney parameters, and 20 corresponding right kidney parameters. From these data, an expert in radionuclide scintigraphy (domain expert) used his experience and published data [15] to estimate five boundary conditions for each parameter: definitely abnormal, probably abnormal, equivocal, probably normal, and definitely normal. A sigmoidlike fit constrained to these five boundary conditions was then generated to create a parameter knowledge library to be used for converting the value of any individual quantitative parameter to a certainty factor (alternative to conditional probability) regarding normality or abnormality. For example, the certainty factor value of 1 indicates that the parameter is definitely abnormal and the value of -1 indicates the parameter is definitely normal; the cutoff certainty factor values between abnormal and equivocal and between normal and equivocal are 0.2 and -0.2 , respectively (Figs. 1A and 1B).

Sixty heuristic rules (IF A THEN B) were extracted from the domain expert to generate the knowledge base for detecting obstruction; 12 of these 60 rules are specifically applied to the baseline study to determine the need for a furosemide administration. Each rule uses the certainty factors describing the degree of abnormality or normality for each parameter that the rule evaluates to generate a certainty factor regarding the need for furosemide to exclude obstruction. These applied rules are chained together by a forward chaining inference engine.

An inference engine is software that selects and executes the rules. The design of the RENEX inference engine follows the MYCIN inference engine by approximating Bayes' theorem to combine the certainty factors generated by the relevant rules to reach a conclusion (combined certainty factor) regarding the need for furosemide. The combined certainty factor can range from "definitely needs furosemide" (1.0) to "definitely does not need furosemide" (-1.0). A meta-rule states that patients whose study has a combined certainty factor in the equivocal range (0.2 to -0.2) should also receive furosemide. A software component called a justification engine was implemented to record the sequence of each rule that was fired and the certainty factor value of all input and output parameters at the time of instantiation to track and justify the logic of the conclusions [4, 5]. The justification engine allows a user to query RENEX to determine the rules and parameter values that justify or explain the software's conclusion regarding the need for furosemide.

To determine the need for furosemide, the entire system was trained using a pilot group of 31 patients (61 kidneys; 10 men and 21 women; mean age, 58.8 ± 17.6 [SD] years). Patients in the training group were obtained from our renal database of patients referred for diuresis renography, and they were specifically selected to include a wide range of responses to develop a complete set of rules for determining the need for furosemide. Based on the decision of an expert panel (majority opinion of three of the authors, two institutions), 34 kidneys required furosemide and 27 kidneys did not. The entire decision support system was fine-tuned and optimized to agree with the expert decisions regarding each kidney in the training set. Processing time per patient is practically instantaneous using a 3.0-GHz PC programmed using IDL (Research Systems, Inc.).

CARTAN

CARTAN used the same pilot group of 31 patients (61 kidneys), the same 47 parameters as used by RENEX, and analyzed each kidney separately. Because there were a large number of parameters (47) relative to the number of kidneys (61), a statistical procedure known as CART (classification and regression trees) was implemented to classify kidneys and patients regarding the need for furosemide [6, 15, 16]. Computations were performed using the "rpart" library in the R freeware statistical analysis package (R Project) to grow and prune decision trees based on the pilot data set [17, 18]. A classification tree was constructed by first identifying a single parameter that determined the best separation between kidneys that required furosemide and kidneys that were not obstructed and did not require furosemide. Next, a second parameter was identified that determined the best separation between the two branches of the previous parameter. This process was continued and combined with cross-validation and pruning back of branches to limit the size of the decision tree and avoid overfitting the data.

To reduce dependency on the training data and to stabilize the algorithm, 1,001 classification trees were constructed by the common statistical technique of bootstrapping the training data [17, 19]. Bootstrap aggregation, or bagging, is a computationally intensive resampling technique that produces a different classification tree for each bootstrap, often involving different dichotomized parameters.

Each of the 1,001 classification trees was subsequently applied to the prospective set of 102 patients (200 kidneys) to predict the need for furosemide. The results of each prediction algorithm (furosemide needed or not needed) of each bootstrap sample were tallied, and the conclusion was based on a simple majority vote regarding the need for furosemide for each of the kidneys in the prospective group.

Determination of the Need for Furosemide in Clinical Practice

To determine the percentage of patients referred for suspected obstruction who required and who did not require furosemide, a retrospective review of renal scans from January 1994 through August 2002 was conducted. A total of 711 MAG3 renal scanning examinations were performed for suspected obstruction between these dates; obstruction was excluded by the baseline scan alone in 221 (31%) of those 711 studies and 490 patients (69%) received furosemide. These percentages were used to select patients in the furosemide and no furosemide categories for the prospective study so that the distribution of patients in the prospective study would match the distribution in clinical practice.

Prospective Study Population and Expert Panel Review

The prospective population consisted of 102 patients (200 kidneys; 56 men and 46 women; mean age, 54.6 ± 16.9 years). None of the 102 patients were included in the training set. To obtain a representative sample, 102 patients were randomly selected: 70 patients were randomly selected from a database of patients referred for suspected obstruction who were given furosemide, and 32 additional patients were randomly selected from a database of patients for whom furosemide was considered unnecessary. As previously indicated, the percentage of patients receiving furosemide (69%) was selected to match the need for furosemide observed in clinical practice. Using only data provided on the scans, three experts (three of the authors) from two different institutions independently scored each kidney on the need for furosemide and resolved differences by consensus. Results from RENEX and CARTAN were prospectively compared with the experts' decision and the actual clinical decision to give or withhold furosemide; results for the left kidney were compared with results for the right kidney. In addition, the experts' decision was compared with the clinical decision to give or withhold furosemide.

Statistical Analysis

Statistical analysis was based on the exact form of the McNemar test for correlated proportions [20]. Due to some small sample sizes and generally good coverage, the Wilson score method was used for calculating confidence intervals for proportions [21, 22]. Significance was set at a p value < 0.05 .

Results

The baseline plus furosemide protocol recommended in the 1996 consensus report on diuresis renography appears to be in common use [1]. On a show of hands or keypad response, 50%, 90%, 85%, and 80% of the participants attending the 2005 Educational Symposia Nuclear Medicine course, the 2005 Emory course in Nuclear Medicine, the 2005 Northeast Regional Meeting of the SNM, and a 2005 RSNA educational course, respectively, used the standard protocol as either two separate acquisitions or a continuous acquisition with furosemide administered after approximately 20 minutes.

Regarding the need for furosemide, RENEX agreed with the clinical decision to give furosemide in 97% of patients (68/70), whereas CARTAN agreed in 90% of patients (63/70), $p = 0.03$ (Table 1). Similarly, RENEX agreed with the experts' decision to give furosemide in 98% of patients (65/66), whereas CARTAN agreed with 89% (59/66), $p =$

0.01 (Table 1). When the data were analyzed by kidney rather than by patient, RENEX agreed with the experts' decision to give furosemide in 94% of kidneys (84/89) compared with 82% (73/89) for CARTAN, $p = 0.002$. Sample studies are illustrated in Figures 2 and 3. The experts' decision agreed with the clinical decision to give furosemide in 94% of patients (66/70).

There was no significant difference between RENEX and CARTAN when the clinicians and experts deemed that furosemide was not necessary, $p = 0.18$ and 0.32 , respectively (Table 2). However, when the data were analyzed by individual kidneys (Table 2), CARTAN agreed with the experts' decision in 78% of kidneys (87/111), whereas RENEX agreed in 69% of kidneys (77/111), $p = 0.008$. The experts' decision agreed with the clinical decision to withhold furosemide in 100% (32/32) of patients.

Overall, RENEX agreed with the clinical decision in 89% (91/102) of patients, with the experts' decision in 87% (89/102) of patients, and with the experts' decision in 81% (161/200) of kidneys. CARTAN agreed with the clinical decision in 87% (89/102) of patients, with the experts' decision in 83% (85/102) of patients, and with the experts' decision in 80% (160/200) of kidneys. The experts' decision agreed with the clinical decision to give or withhold furosemide in 96% (98/102) of patients.

The performance of RENEX and CARTAN in the right kidney was compared with the performance in the left kidney. RENEX agreed with experts' decision to administer furosemide in 93% (40/43) of left kidneys and 96% (44/46) of right kidneys, $p =$ not statistically significant (Table 3). Comparable results were obtained for CARTAN (Table 3). Similarly, there was no significant difference between the right and left kidneys for RENEX and CARTAN when the results were analyzed based on the experts' decisions to withhold furosemide (Table 4).

Discussion

Over the past several years, artificial intelligence methods have been investigated as a way to develop tools to improve diagnostic performance. Examples from nuclear medicine include neural networks [23–25] and case-based reasoning techniques [26] to provide computer-assisted diagnosis of planar and SPECT myocardial perfusion studies. In knowledge-based expert systems, a knowledge base of heuristic rules is obtained from human experts capturing how they make their interpretations. These rules are usually expressed in the form of IF A THEN B expressions.

One major advantage of a knowledge-based expert system is that the system does not require the same large numbers of studies as are required to develop neural networks or case-based reasoning approaches. A second advantage, especially from a learning perspective, is that it is possible to query a knowledge-based system to learn the rules that led to a specific conclusion. For example, RENEX disagreed with the experts regarding the need for furosemide in the right kidney (Fig. 3), but it is possible to query RENEX to determine the reasons for its decision. The primary reason RENEX concluded furosemide was needed was that the half-time ($T_{1/2}$) was abnormal (20.8 minutes); RENEX recognized that there was moderate evidence that furosemide was not needed because the postvoid to maximum ratio was 0.2 (upper limits of the normal value of 0.16) [15], but RENEX gave greater weight to the abnormal $T_{1/2}$. The experts were not queried but clearly gave greater weight to the images and the low postvoid to maximum ratio. RENEX can be improved by comparing results and “reasoning” with expert decisions and adding new rules or adjusting the weighting factors. In this case, RENEX should be tested giving greater weight to the postvoid to maximum count ratio than the $T_{1/2}$.

We chose to implement a purely statistical method (CARTAN) for the second decision support approach because the combined approach of classification and regression tree (CART) analysis with bootstrapping could handle a relatively large number of variables and arrive at a prediction algorithm with a relatively small training set.

CART analysis, but not bagging, has been used to help distinguish between healthy and diseased kidneys using MR data that include measurements of renal arterial blood flow and parenchymal perfusion [27]. To our knowledge, this is the first time CART analysis and bagging have been applied to a problem in nuclear medicine. CART agreed with the expert decisions illustrated in Figure 3, but CART cannot provide the rationale for these decisions. However, CART can use a large sample of test patients to determine the parameters and parameter cutoff values that provide the best separation in distinguishing when furosemide is needed or not needed or whether a kidney is obstructed. These parameters and cutoff values can be used to guide clinicians and RENEX in determining the need for furosemide and the presence or absence of obstruction.

Our long-term goal is to develop decision support systems to determine if a kidney is obstructed. The dual statistical and knowledge-based approaches have provided complementary information and helped target features that can be transferred or emphasized to improve the overall performance of both systems. At present, data are insufficient to determine if one approach is inherently superior to the other. Analysis of the baseline study to determine the need for furosemide may be useful for the 50–90% of institutions that use a baseline plus furosemide protocol. These decision support systems have the potential to serve as a second opinion to support a clinician's preliminary decision regarding the need for furosemide. This study has also provided important data and direction for developing decision support systems to diagnose obstruction.

The focus of this research effort is to develop statistical and knowledge-based decision support systems that will have diagnostic performance in diuresis renography comparable to that of experts. This is not a trivial task; however, it could be argued that a better goal would be to develop decision support systems that use the clinical outcome as the gold standard rather than expert reviewers. This is an attractive goal, but it is confounded by the fact that an obstructed or nonobstructed scan interpretation has a major impact on the clinical outcome. Consequently, this gold standard becomes biased.

Additional problems of data analysis can be illustrated by a patient who had a pyeloplasty to relieve obstruction 6–24 months after a diuresis renography scan was interpreted as no obstruction. In this example, did the scan miss the obstruction, was the study interpreted incorrectly, or did the patient become obstructed 6–24 months after the scan? Using patient outcome as a gold standard can be an important goal, but interpretation of the results is not straightforward. Moreover, regardless of the type of study, clinical practice would be improved if radiologists can provide an interpretation equivalent to that of expert reviewers.

The study has several limitations. The training set was relatively small; this is probably more of a disadvantage for a statistical system such as CARTAN than for a knowledge-based system, but both of these systems could potentially be improved by increasing the size of the training set. This study addressed the diuresis renography protocol recommended by the international consensus report in which baseline data are obtained, followed by the administration of furosemide and an additional period of imaging [1]. There are other protocols in which furosemide is given 15 minutes before the radiopharmaceutical, at the same time as the radiopharmaceutical, or 5 to 10 minutes later [1, 28, 29]. Obviously, the systems we describe do not apply to these protocols. Finally, QuantEM 2.0 cannot detect and correct for patient motion, and at this time the software cannot distinguish between diffuse

retention with slow washout due to impaired function and focal pelvic retention with slow washout due to possible obstruction. Algorithms to detect and correct for motion and to distinguish between diffuse retention in a kidney and retention in a dilated renal collecting system need to be designed, implemented, and tested.

In conclusion, RENEX performed better than CARTAN when clinicians and experts determined that furosemide was necessary to evaluate obstruction. When furosemide was not required, CARTAN performed better than RENEX for cases in which the results were analyzed by individual kidney; there was no difference when the results were analyzed by patient. Use of RENEX or CARTAN as decision support tools in the baseline plus furosemide protocol has the potential to help avoid unnecessary imaging; minimize patient inconvenience; and reduce costs by saving technologist, camera, computer, and physician time. Notably, the results obtained from this study will assist in the more complex task of developing decision support systems for diuresis renography to evaluate obstruction.

Acknowledgments

We thank Russell D. Folks, Angela Hunsche, and Meghna Krishnan for their assistance in collecting and archiving the data used in this study.

Supported by grant ROI LMN07595 from the Library of Medicine.

References

1. 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; 37:1872–1876. [PubMed: 8917195]
2. Kuyvenhoven J, Piepsz A, Ham H. When could the administration of furosemide be avoided? *Clin Nucl Med.* 2003; 28:732–737. [PubMed: 12972994]
3. IMV Medical Information Division. 2003 nuclear medicine census market summary report. Des Plaines, IL: IMV Limited; 2003. p. 7-11.
4. Garcia EV, Halkar R, Folks R, Krishna M, Taylor A. RENEX: an expert system for the interpretation of Tc-99m MAG3 scans to detect renal obstruction. *J Nucl Med.* 2005; 46:205P–206P.
5. Garcia EV, Taylor A, Halkar R, et al. RENEX: an expert system for the interpretation of Tc-99m MAG3 scans to detect renal obstruction. *J Nucl Med.* 2006; 47:320–329. [PubMed: 16455639]
6. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. Classification and regression trees. Belmont, CA: Wadsworth; 1984.
7. Taylor A Jr, Corrigan PL, Galt J, et al. Measuring technetium-99m-MAG3 clearance with an improved camera-based method. *J Nucl Med.* 1995; 36:1689–1695. [PubMed: 7658232]
8. Taylor A, Manatunga A, Morton K, et al. Multi-center trial validation of a camera-based method to measure Tc-99m mercaptoacetyl triglycine, or Tc-99m MAG3, clearance. *Radiology.* 1997; 204:47–54. [PubMed: 9205222]
9. Taylor A, Lewis C, Giacometti A, et al. Improved formulas for the estimation of renal depth in adults. *J Nucl Med.* 1993; 34:1766–1769. [PubMed: 8410296]
10. Taylor A. Formulas to estimate renal depth in adults. *J Nucl Med.* 1994; 35:2054–2055.
11. Shortliffe, EH., et al. Computer-based medical consultations: MYCIN. Amsterdam, The Netherlands: Elsevier Scientific Publishing; 1976. p. 264
12. Garcia EV, Cooke CD, Folks RD, et al. Diagnostic performance of an expert system for the interpretation of myocardial perfusion SPECT studies. *J Nucl Med.* 2001; 42:1185–1191. [PubMed: 11483678]
13. Ezquerria N, Mullick R, Cooke D, Krawczynska E, Garcia E. PERFEX: an expert system for interpreting 3D myocardial perfusion. *Expert Systems with Applications.* 1993; 6:459–468.

14. Esteves FP, Halkar RK, Issa MM, Grant S, Taylor A. Comparison of camera-based Tc-99m MAG3 and 24-hour creatinine clearances for evaluation of kidney function. *AJR*. 2006; 187:833, W316–W319. [Web].
15. Esteves F, Taylor A, Manatunga A, Folks R, Krishnan M, Garcia E. Normal values for camera-based ^{99m}Tc-MAG3 clearance, MAG3 curve parameters, excretory parameters, and residual urine volume. *AJR*. 2006; 187:1563, W1–W8. [Web].
16. Venables, WN.; Ripley, BD. *Modern applied statistics with S-plus*. 3rd ed.. New York, NY: Springer-Verlag; 1999. p. 303-310.
17. Hastie, T.; Tibshirani, R.; Friedman, J. *The elements of statistical learning: data mining, inference, and prediction*. New York, NY: Springer; 2001. p. 246-250.
18. R Project Web site. [Accessed January 19, 2007] Comprehensive R Archive Network (CRAN). Available at www.r-project.org.
19. Efron, B.; Tibshirani, R. *An introduction to the bootstrap*. London, UK: Chapman and Hall; 1993.
20. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*. 1947; 12:153–157. [PubMed: 20254758]
21. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med*. 1998; 17:857–872. [PubMed: 9595616]
22. Agresti A, Coull BA. Approximate is better than “exact” for interval estimation of binomial proportions. *Am Statistician*. 1998; 52:119–126.
23. Fujita H, Katafuchi T, Uehara T, Nishimura T. Application of neural network to computer-aided diagnosis of coronary artery disease in myocardial SPECT bull’s-eye images. *J Nucl Med*. 1992; 33:272–276. [PubMed: 1732455]
24. Porenta G, Dorffner G, Kundrat S, Petta P, Duit-Schedlmayer J, Sochor H. Automated interpretation of planar thallium-201-dipyridamole stress-redistribution scintigrams using artificial neural networks. *J Nucl Med*. 1994; 35:2041–2047. [PubMed: 7989989]
25. Hamilton D, Riley PJ, Miola UJ, Amro AA. A feed forward neural network for classification of bull’s-eye myocardial perfusion images. *Eur J Nucl Med*. 1995; 22:108–115. [PubMed: 7758496]
26. Haddad M, Adlassnig KP, Porenta G. Feasibility analysis of a case-based reasoning system for automated detection of coronary heart disease from myocardial scintigrams. *Artif Intell Med*. 1997; 9:61–78. [PubMed: 9021059]
27. Michaely HJ, Schoenberg SO, Carina I, Dikow R, Bock M, Guenther M. Renal disease: value of functional magnetic resonance imaging with flow and perfusion measurements. *Invest Radiol*. 2004; 39:698–705. [PubMed: 15486531]
28. Sfakianakis GN, Cohen DJ, Braunstein RH, et al. MAG3-F₀ scintigraphy in decision making for emergency intervention in renal colic after helical CT positive for a urolith. *J Nucl Med*. 2000; 41:1813–1822. [PubMed: 11079488]
29. Wong DC, Rossleight MA, Farnsworth RH. F + 0 diuresis renography in infants and children. *J Nucl Med*. 1999; 40:1805–1811. [PubMed: 10565774]

Fig. 1. Knowledge library was generated for each parameter to convert any given parameter value into certainty factor

A, Graph shows parameter knowledge library to convert any time to half maximum count value for left kidney (whole-kidney region of interest [ROI]) to certainty factor value. Boundary values (see Subjects and Methods section for more detail) were 8, 11, 13, 15, and 20 minutes, respectively.

B, Graph shows parameter knowledge library to convert any 20 minute to maximum count ratio for left kidney (cortical ROI) to certainty factor value. Boundary values were 0.17, 0.27, 0.32, 0.43, and 0.92, respectively.

Fig. 2. 50-year-old woman referred for mercaptoacetyl triglycine (MAG3) scanning because of suspected obstruction

A, QuantEM (developed at Emory University, licensed to GE Healthcare) review of baseline scan shows no infiltration. MAG3 clearance is reduced. Right kidney shows prompt uptake and excretion and is not obstructed. Left kidney shows reduction in relative function (18%). There was uniform uptake in left kidney with slow washout, but, importantly, there was no retention of tracer in collecting system. Clinically, study was interpreted as not obstructed and furosemide was not administered. RENEX, CARTAN, and experts agreed with this decision. Tmax = time to maximum counts, T_{1/2} = half-time, 20 min/max = 20-minute-to-maximum-count ratio, k = kilo.

B, Expanded QuantEM review display shows patient values for MAG3 clearance; residual urine volume; percent relative uptake; and time to maximum counts (Tmax), 20-minute-to-maximum-count ratio (20/max), half-time ($T_{1/2}$), and postvoid to maximum (postvoid/max) ratios for whole-kidney and cortical regions of interest as well as normal ranges for each of these values. Expanded review page also shows enlarged parenchymal image obtained at 2–3 minutes, enlarged display of 19- to 20-minute image, and quality control images showing preinjection and postinjection syringe counts and time of bolus arrival in kidneys. Postvoid to maximum count ratio of left kidney is slightly elevated at 23%, but this ratio tells clinician that almost 80% of maximal activity has washed out of left kidney by conclusion of study. Such high washout percentage is strong evidence against obstruction.

Fig. 3. 40-year-old man referred for mercaptoacetyl triglycine (MAG3) scanning because of suspected obstruction

A, QuantEM (developed at Emory University, licensed to GE Healthcare) review of baseline scan shows no infiltration. Camera-based MAG3 clearance is normal. Left kidney shows rapid uptake and washout of tracer in sequential 2-minute images (*center*). Relative uptake of left kidney is 56%, whole-kidney half-time ($T_{1/2}$) is 10.8 minutes, and postvoid to maximum count ratio (postvoid/max) is normal (**B**) excluding obstruction. Relative function of right kidney was 46%, and right kidney showed dilatation of renal pelvis, which can be appreciated on 2-minute image as well as on enlarged 2- to 3-minute image (**B**). $T_{1/2}$ for whole-kidney region of interest (ROI) was prolonged at 20.3 minutes and 20-minute-to-

maximum-count ratio (20/max) was elevated at 0.58. However, visually pelvis empties after voiding and postvoid/max was 0.20 (**B**). Although postvoid/max count ratio is slightly elevated (0.16 is the upper limit of normal [15]), experts' and clinical decisions were that furosemide was not indicated. CARTAN agreed with experts, but RENEX concluded that furosemide was needed to exclude obstruction of right kidney (see Discussion). Tmax = time to maximum count ratio, 20/max = 20-minute-to-maximum-count ratio, k = kilo. **B**, Expanded QuantEM review display shows patient values for MAG3 clearance; residual urine volume; percent relative uptake; and time to maximum (Tmax), 20-minute-to-maximum-count ratio (20/max), half-time ($T_{1/2}$), and postvoid to maximum count (postvoid/max) ratios for whole-kidney and cortical ROIs as well as normal ranges for each of these values. Expanded review page also shows enlarged parenchymal image obtained at 2–3 minutes, enlarged display of 19- to 20-minute image, and quality control images showing preinjection and postinjection syringe counts and time of bolus arrival in kidneys. In particular, enlarged 2- to 3-minute image shows dilated renal pelvis, and postvoid/max ratio for right kidney is slightly elevated.

TABLE 1
Performance of RENEX and CARTAN Compared with Expert Consensus and Clinical Decision to Administer Furosemide

Decision Source	No. Requiring Furosemide	RENEX	Correct Proportion (95% CI)	CARTAN	Correct Proportion (95% CI)	p
Clinical decision (patients)	70	68	0.97 (0.90–0.99)	63	0.90 (0.81–0.95)	0.03
Consensus (patients)	66	65	0.98 (0.92–1.00)	59	0.89 (0.80–0.95)	0.01
Consensus (kidneys)	89	84	0.94 (0.88–0.98)	73	0.82 (0.73–0.89)	0.002

Note—CI = confidence interval.

TABLE 2
Performance of RENEX and CARTAN Compared with Expert Consensus and Clinical Decision Furosemide Was Unnecessary

Decision Source	No. Not Requiring Furosemide	RENEX	Correct Proportion (95% CI)	CARTAN	Correct Proportion (95% CI)	<i>p</i>
Clinical decision (patients)	32	23	0.72 (0.55–0.84)	26	0.81 (0.65–0.91)	0.18
Consensus (patients)	36	24	0.67 (0.50–0.80)	26	0.72 (0.56–0.84)	0.32
Consensus (kidneys)	111	77	0.69 (0.60–0.77)	87	0.78 (0.70–0.85)	0.008

Note—CI = confidence interval.

TABLE 3
Performance of RENEX and CARTAN Compared with Expert Consensus to Administer Furosemide for Right and Left Kidneys

Kidney	No. Not Requiring Furosemide	RENEX	Correct Proportion (95% CI)	CARTAN	Correct Proportion (95% CI)
Left	43	40	0.93 (0.81–0.98)	35	0.81 (0.67–0.90)
Right	46	44	0.96 (0.85–0.99)	38	0.83 (0.69–0.91)

TABLE 4

Performance of RENEX and CARTAN Compared with Expert Consensus Furosemide Was Not Necessary for Right and Left Kidneys

Kidney	No. Not Requiring Furosemide	RENEX	Correct Proportion (95% CI)	CARTAN	Correct Proportion (95% CI)
Left	55	39	0.71 (0.58–0.81)	45	0.82 (0.70–0.90)
Right	56	38	0.68 (0.55–0.79)	42	0.80 (0.62–0.84)