Expert systems for diuresis renography
Andrew Taylor, MD
Emory University, Atlanta, GA, USA
Collaborators

Ernest V. Garcia, PhD
Amita Manatunga, PhD
Jose Binonngo, PhD
Jieqiong Bao, MS
Russell Folks, CNMT
Daya Manatunga, MS
Raghuveer Halkar, MD
Eva Dubovsky, MD, PhD
Radiologists interpret the majority of scans in US

Anatomy vs physiology

3 months nuclear medicine training for radiologists

36 months for nuclear medicine physicians

Nuclear medicine is a small part of radiology

Renal nuclear medicine is a small part of Nucl Med
THE PROBLEM

How is the general radiologist to acquire the experience and expertise to interpret these renal studies?
Urologists have problems too.

Quote from my brother-in-law (a radiologist in Los Angeles)

“My (referring) urologist doesn’t care about my interpretation of furosemide renal scans. All he wants to know is the T1/2. He says if the T1/2 is greater than 20 minutes, he operates.”
SOLUTION?

CONSENSUS REPORTS
Consensus Reports

Diuresis renography (1996): O’Reilly et al. JNM
Clearance measurements (1996): Blaufox et al. JNM
ACE inhibition renography (1996): Taylor et al. JNM
QC, renogram measurements (1999): Prigent et al. Seminars
Transplants, technical aspects (1999). Dubovsky et al. Seminars
Renal transit times (2008). Durand et al. Seminars
Consensus Reports

Survey of full time nuclear medicine practitioners in Britain.

Only 49% of physicians were even aware a guideline on clearance measurements existed.

Cosgriff and Stevens. Alasbimn Journal 2001
Another solution?

Software to interpret renal scans
Expert (Decision Support) Systems

- **Neural networks**
  - Repeated recognition trials to train net and predict output
  - Need many cases
  - No rationale for the decision

- **Case-based reasoning**
  - Algorithm searches a library of patient cases for best match
  - Need many cases
  - No rationale for the decision
Expert (Decision Support) Systems

• Predictive statistical models
  – Step-wise regression, backward elimination, clustering
  – Determine parameters and weighting factors associated with scan interpretation
  – Incorporate this information into a predictive algorithm.
  – Many cases

• Heuristic models (knowledge based expert systems)
  – If, then statements (rules)
  – Certainty factors
  – Fewer cases
  – Provides rationale
RENEX
STATEX
Reason for referral*

1. Obstruction (49%)
2. Renovascular hypertension (17%)
3. Transplant evaluation (15%)
4. Renal function (relative/absolute) 14%
5. Other (4%)
7. Evaluation of pediatric patients (0%)

*1415 renal scans between 1994 and 2002
Goals for expert (decision support) system

• Provide renal scan analysis and interpretation
• Minimize renal scan acquisition and processing errors
• Reduce diagnostic errors
• Save physician time
• Train residents (nuclear med, radiologists, urologists)
• Develop a template that can be applied to other problems
• Assist, not replace, physicians
“I would rather have a good technologist than a good computer program.”
Another comment from a reviewer

“The physician has to do something to justify getting paid.”
Protocol for Diuresis Renography

1. One acquisition, furosemide given at 20 min (T + 20)*
2. Baseline scan; 2nd acquisition with furosemide
3. One acquisition, furosemide given at T = -15
4. One acquisition, furosemide given at T = 0
5. One acquisition, furosemide given at 10 min (T +10)

QuantEM\textsuperscript{TM} 2.0 acquisition and processing

- IDL\textsuperscript{1}; runs on standard Windows PC
- 2 sec frames for 5 min, 10 sec frames for 19 min
- Camera based MAG3 clearance (1-2.5 min)\textsuperscript{2}
- Automated whole kidney, cortical, background ROIs
- Post-void image of kidneys, bladder
- Separate baseline and furosemide acquisitions
- Forwards 47 baseline parameters to DSS
- Forwards 19 post-furosemide parameters to DSS

\textsuperscript{1}Interactive Data Language, RSI, Boulder, CO
\textsuperscript{2}Taylor et al. Radiology 1997;2004:47
Normal values

- Camera-based MAG3 clearance
- Whole kidney/cortical curve parameters
- Excretory parameters
- Residual urine volume
- 106 potential renal donors

### Normal Values Using ROIs over the Whole Kidney

<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>5th percentile</th>
<th>95th percentile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tmax, right kidney (min) &amp;</td>
<td>M</td>
<td>44</td>
<td>3.57</td>
<td>2.1</td>
<td>2.1</td>
<td>2.3</td>
<td>6.3</td>
<td>15.3</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>62</td>
<td>4.35</td>
<td>2.7</td>
<td>2.3</td>
<td>2.3</td>
<td>9.8</td>
<td>16.3</td>
</tr>
<tr>
<td>Tmax, left kidney (min) &amp;</td>
<td>M</td>
<td>44</td>
<td>3.16</td>
<td>1.0</td>
<td>2.1</td>
<td>2.1</td>
<td>5.8</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>62</td>
<td>3.72</td>
<td>1.7</td>
<td>2.3</td>
<td>2.3</td>
<td>6.8</td>
<td>11.3</td>
</tr>
<tr>
<td>T ½, right kidney (min) &amp;</td>
<td>M</td>
<td>44</td>
<td>5.64</td>
<td>2.3</td>
<td>2.0</td>
<td>3.3</td>
<td>8.3</td>
<td>16.5</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>62</td>
<td>8.29</td>
<td>8.4</td>
<td>3.5</td>
<td>4.0</td>
<td>17.0</td>
<td>50.0</td>
</tr>
<tr>
<td>T ½, left kidney (min)</td>
<td>M</td>
<td>44</td>
<td>5.36</td>
<td>1.4</td>
<td>3.0</td>
<td>3.5</td>
<td>7.5</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>62</td>
<td>6.26</td>
<td>2.8</td>
<td>3.0</td>
<td>3.7</td>
<td>10.5</td>
<td>18.3</td>
</tr>
<tr>
<td>20 min/max ratio (right)</td>
<td></td>
<td>106</td>
<td>0.24</td>
<td>0.14</td>
<td>0.11</td>
<td>0.12</td>
<td>0.54</td>
<td>0.96</td>
</tr>
<tr>
<td>20 min/max ratio (left)</td>
<td></td>
<td>106</td>
<td>0.22</td>
<td>0.08</td>
<td>0.11</td>
<td>0.13</td>
<td>0.35</td>
<td>0.61</td>
</tr>
<tr>
<td>20 min/2-3 min ratio (right)</td>
<td></td>
<td>106</td>
<td>0.24</td>
<td>0.19</td>
<td>0.09</td>
<td>0.11</td>
<td>0.64</td>
<td>1.35</td>
</tr>
<tr>
<td>20 min/2-3 min ratio (left)</td>
<td></td>
<td>106</td>
<td>0.20</td>
<td>0.09</td>
<td>0.09</td>
<td>0.11</td>
<td>0.34</td>
<td>0.62</td>
</tr>
</tbody>
</table>

*There is no significant difference between younger (< 40 years) and older (> 40 years) adults.

&There is a significant difference (p<0.05) between males and females.

Std. Dev. = standard deviation
<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>5&lt;sup&gt;th&lt;/sup&gt; percentile</th>
<th>95&lt;sup&gt;th&lt;/sup&gt; percentile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tmax, right kidney (min) &amp;</strong></td>
<td>M</td>
<td>44</td>
<td>3.57</td>
<td>2.1</td>
<td>2.1</td>
<td>2.3</td>
<td>6.3</td>
<td>15.3</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>62</td>
<td>4.35</td>
<td>2.7</td>
<td>2.3</td>
<td>2.3</td>
<td>9.8</td>
<td>16.3</td>
</tr>
<tr>
<td><strong>Tmax, left kidney (min) &amp;</strong></td>
<td>M</td>
<td>44</td>
<td>3.16</td>
<td>1.0</td>
<td>2.1</td>
<td>2.1</td>
<td>5.8</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>62</td>
<td>3.72</td>
<td>1.7</td>
<td>2.3</td>
<td>2.3</td>
<td>6.8</td>
<td>11.3</td>
</tr>
<tr>
<td><strong>T ½, right kidney (min) &amp;</strong></td>
<td>M</td>
<td>44</td>
<td>5.64</td>
<td>2.3</td>
<td>2.0</td>
<td>3.3</td>
<td>8.3</td>
<td>16.5</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>62</td>
<td>8.29</td>
<td>8.4</td>
<td>3.5</td>
<td>4.0</td>
<td>17.0</td>
<td>50.0</td>
</tr>
<tr>
<td><strong>T ½, left kidney (min)</strong></td>
<td>M</td>
<td>44</td>
<td>5.36</td>
<td>1.4</td>
<td>3.0</td>
<td>3.5</td>
<td>7.5</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>62</td>
<td>6.26</td>
<td>2.8</td>
<td>3.0</td>
<td>3.7</td>
<td>10.5</td>
<td>18.3</td>
</tr>
<tr>
<td><strong>20 min/max count ratio (right)</strong></td>
<td>106</td>
<td></td>
<td>0.24</td>
<td>0.14</td>
<td>0.11</td>
<td>0.12</td>
<td>0.54</td>
<td>0.96</td>
</tr>
<tr>
<td><strong>20 min/max count ratio (left)</strong></td>
<td>106</td>
<td></td>
<td>0.22</td>
<td>0.08</td>
<td>0.11</td>
<td>0.13</td>
<td>0.35</td>
<td>0.61</td>
</tr>
<tr>
<td><strong>20 min/2-3 min count ratio (right)</strong></td>
<td>106</td>
<td></td>
<td>0.24</td>
<td>0.19</td>
<td>0.09</td>
<td>0.11</td>
<td>0.64</td>
<td>1.35</td>
</tr>
<tr>
<td><strong>20 min/2-3 min count ratio (left)</strong></td>
<td>106</td>
<td></td>
<td>0.20</td>
<td>0.09</td>
<td>0.09</td>
<td>0.11</td>
<td>0.34</td>
<td>0.62</td>
</tr>
</tbody>
</table>

*There is no significant difference between younger (< 40 years) and older (> 40 years) adults.

&There is a significant difference (p<0.05) between males and females.

Std. Dev. = standard deviation
Camera based MAG3 clearance vs creatinine clearance

Halkar et al. Urology 2007;69:426
Do we need to give furosemide? (Can obstruction be excluded by the baseline study?)

Compared to consensus of 3 experts and clinical decision as gold standard.
We have to learn to walk before we can run.
1. When can obstruction be excluded without giving furosemide

2. What are the key parameters

3. Training set (64 patients/124 kidneys)

4. Validation set (80 patients/157 kidneys)

Binongo JNG et al. Use of classification and regression trees in diuresis renography. Acad Radiol 2007;14:306-311
STATEX

1. $T \frac{1}{2}$ (cortical/whole kidney)
2. $T_{max}$ (cortical/whole kidney)
3. Relative uptake
4. Post-void/max ratio
5. MAG3 clearance
6. 20 min/max ratio (cortical/whole kidney)
7. 20 min/2-3 min ratio (cortical/whole kidney)
8. Renal size (pixel area)
Lasix needed: 29/30
1/30 (3%)

Lasix needed: 5/5
0/5 (0%)

No lasix: 34/34
0/34 (0%)

No lasix: 2/2
0/2 (0%)

No lasix: 4/4
0/2 (0%)

Lasix needed: 3/4
1/4 (25%)
Misclassification Rate

Training Set:  2.5%

Validation Set: 15.6%
Bootstrap aggregation (bagging)

Training set

Bootstrap Set 1

Bootstrap Set 2

Bootstrap Set 1001
## Misclassification rates with original tree and bagging ($k = 1001$)

<table>
<thead>
<tr>
<th>Kidney</th>
<th>Number</th>
<th>Original Tree</th>
<th>Bagging</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right</td>
<td>64</td>
<td>15.6%*</td>
<td>10.9%*</td>
</tr>
<tr>
<td>Left</td>
<td>60</td>
<td>8.3%</td>
<td>10.0%</td>
</tr>
</tbody>
</table>

*P = 0.03
Lasix needed: 29/30
1/30 (3%)

≥ 0.38

post-void/max ratio

Lasix needed: 5/5
0/5 (0%)

< 180
cortical 20min/max ratio

≥ 86%

Lasix needed: 3/4
1/4 (25%)

≥ 68%

No lasix: 4/4
0/2 (0%)

≥ 0.51

No lasix: 2/2
0/2 (0%)

≥ 180

MAG3 clearance

Lasix needed: 5/5
0/5 (0%)

< 68%

No lasix: 34/34
0/34 (0%)

< 0.38

< 0.51

< 86%

Level 0

Level 1

Relative uptake

Level 2

cortical 20min/max ratio

Level 3

Relative uptake

Level 4
Percent of 1001 bootstrapped trees from the training set containing one or more levels

<table>
<thead>
<tr>
<th>Level</th>
<th>Right Kidney</th>
<th>Left Kidney</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.7%</td>
<td>9.2%</td>
</tr>
<tr>
<td>2</td>
<td>25.7%</td>
<td>41.6%</td>
</tr>
<tr>
<td>3</td>
<td>44.8%</td>
<td>37.7%</td>
</tr>
<tr>
<td>4</td>
<td>23.1%</td>
<td>10.5%</td>
</tr>
<tr>
<td>5</td>
<td>2.7%</td>
<td>1.0%</td>
</tr>
<tr>
<td>6</td>
<td>0.0%</td>
<td>0.1%</td>
</tr>
</tbody>
</table>
1. T ½ (cortical/whole kidney)
2. Tmax (cortical/whole kidney)
3. Relative uptake
4. Post-void/max ratio
5. MAG3 clearance
6. 20 min/max ratio (cortical/whole kidney)
7. 20 min/2-3 min ratio (cortical/whole kidney)
8. Renal size (pixel area)
Most frequent kidney parameters for excluding obstruction (training set, level 1)

<table>
<thead>
<tr>
<th>Kidney Parameter</th>
<th>Right Kidney</th>
<th>Left Kidney</th>
</tr>
</thead>
<tbody>
<tr>
<td>postvoid/max ratio</td>
<td>29.1%</td>
<td>83.0%</td>
</tr>
<tr>
<td>cortical 20 min/max ratio</td>
<td>32.1%</td>
<td>3.6%</td>
</tr>
<tr>
<td>postvoid/1-2 min ratio</td>
<td>13.3%</td>
<td>6.3%</td>
</tr>
<tr>
<td>Kidney 20 min/max ratio</td>
<td>10.8%</td>
<td>1.0%</td>
</tr>
</tbody>
</table>
RENEX has to determine if each parameter is normal, equivocal or abnormal.
CERTAINTY FACTOR:
Degree of Belief in a Hypothesis

Hypothesis: Clinical finding is abnormal

Definitely Normal

Indeterminate

Definitely Abnormal

-1.0

-0.2

+0.2

1.0 CF

Sigmoid Transformation from Qt Parameters to CF values: Boundary Conditions

- Definitely Abnormal (CF = 1.0)
- Definitely Normal (CF = -1.0)
- Normal (CF ≤ -0.2)
- Abnormal (CF ≥ 0.2)
- Equivocal (CF = 0)

Quantitative Parameter Value

CF Value

0

-1.0

-0.5

0

0.5

1.0

CF Value

Quantitative Parameter Value

+1.0

Definitely Abnormal (CF = 1.0)

Definitely Normal (CF = -1.0)

Equivocal (CF = 0)

Abnormal (CF ≥ 0.2)

Normal (CF ≤ -0.2)
Conversion of Four Input Parameters

- LK MAG3 Clearance
- Lasix Prevoid/Baseline max
- Lasix Postvoid/Baseline 1-2min
- Lasix Pelvis Time to Half Peak
PATIENT MAG3 STUDY ACQUISITION

PROCESSING QUANTIFICATION: QuantEM 2.0

Image file

Patient Qt Parameters

CONVERT TO CF

Parameter Knowledge File

SIGMOID FIT of 5 BOUNDARY CONDITIONS

Quantitative PARAMETER INPUT LIST

NORMAL DATABASE & CRITERIA

RENEX: Information Flow
If the time to half peak of the left kidney pelvis after furosemide is abnormal (CF > 0.2), then there is a strong positive evidence (CF=.4) that the left kidney is obstructed.

http://www.med.emory.edu/decisionsupportsystems.com
Inference Engine

Combines certainty factors using logic (MYCIN) which incrementally adjusts the certainty factor of a diagnosis as evidence becomes available.

CF Modification as Rules are Fired
Justification Engine

- Justification engine keeps track of the sequence of rules that are fired to reach a conclusion
- Interpreter can request justification by clicking on a specific conclusion
- Justification process reports *English translation* of all concatenated rules used to reach the conclusion
**RENEX: Information Flow**

- **NORMAL DATABASE & CRITERIA**

- **PATIENT MAG3 STUDY ACQUISITION**
  - Quantitative PARAMETER INPUT LIST
  - SIGMOID FIT of 5 BOUNDARY CONDITIONS

- **PROCESSING QUANTIFICATION: QuantEM 2.0**
  - Patient Qt Parameters

- **CONVERT TO CF**

- **KNOWLEDGE BASE**
  - RENEX INFEERENCE ENGINE (60 rules)
  - JUSTIFICATION ENGINE
  - Param. CF Value

- **CF VALUES FOR**
  - Left Kidney Obstruction
  - Right Kidney Obstruction

**Parameter Knowledge File**
STATEX vs RENEX

- Is furosemide needed?
- Same 31 patient training set
- Same data available to both systems
- CART with bagging
- Prospectively applied to 102 patients (200 kidneys)

<table>
<thead>
<tr>
<th>EXPERTS</th>
<th>Number of kidneys</th>
<th>RENEX % correct</th>
<th>STATEX % correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not obstructed</td>
<td>111</td>
<td>69%*</td>
<td>78%*</td>
</tr>
<tr>
<td>No furosemide</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Give furosemide to</td>
<td>89</td>
<td>94%*</td>
<td>82%*</td>
</tr>
<tr>
<td>evaluate obstruction</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* P < 0.05
To Err Is Human
Building a safer health system
Institute of Medicine

“98,000 people die every year in the United States due to medical errors.”
“To err is human”

• Pounds entered for kilograms (250 kg patient)
• Camera started late
• Background ROI outside field of view
• Poor renal function (cortical ROI values)
• Unlikely or unreliable values, patient motion

Data input errors
Acquisition errors
Processing errors
**METHODS**

*User input*

- Change the prompt
- Highlight the value

<table>
<thead>
<tr>
<th>Patient Height:</th>
<th>163.000 (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Patient Weight</em>:</td>
<td>158.000 (kg)</td>
</tr>
<tr>
<td>Patient Age:</td>
<td>67 (yrs)</td>
</tr>
<tr>
<td>Select patient gender:</td>
<td>female</td>
</tr>
<tr>
<td>Dose Injected:</td>
<td>9.98000 (mCi)</td>
</tr>
<tr>
<td>Dose Counted on Camera:</td>
<td>1.84000 (mCi)</td>
</tr>
<tr>
<td>Voided urine volume:</td>
<td>112 (ml)</td>
</tr>
<tr>
<td>Time of initial void (24hr clock):</td>
<td>hr: 10 min: 23</td>
</tr>
<tr>
<td>Time of final void (24hr clock):</td>
<td>hr: 11 min: 9</td>
</tr>
</tbody>
</table>
Development and prospective evaluation of a automated software system for quality control of quantitative $^{99\text{m}}$Tc MAG3 studies

- IDL; runs on a Windows PC
- Monitors user input data
- Flags unusual or impossible values
- Allows user to change/verify unusual values
- Monitors reported values
- Stores results in XML file

Now, for the hard part.

How well does RENEX work for obstruction?
Patient population: 32 patients who received furosemide to evaluate obstruction

Gold standard: 3 experts blindly scored each kidney as obstructed, equivocal or non-obstructed and resolved differences by consensus

Characteristics: 41 non-obstructed, 9 equivocal, 13 obstructed kidneys

Garcia et al. N Nucl Med 2006:47;320
<table>
<thead>
<tr>
<th></th>
<th>Human Experts</th>
<th>Obstructed Kidneys</th>
<th>Equivocal</th>
<th>Not Obstructed</th>
</tr>
</thead>
<tbody>
<tr>
<td>RENEX</td>
<td>Obstructed Kidneys</td>
<td>12</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Equivocal</td>
<td>0</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Not Obstructed</td>
<td>1</td>
<td>0</td>
<td>38</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>92% (12/13)</td>
<td>78% (7/9)</td>
<td>93% (38/41)</td>
<td></td>
</tr>
</tbody>
</table>

Garcia et al. N Nucl Med 2006:47;320
Testing the justification engine

- Prospective group of 60 patients (117 kidneys)
- Expert evaluated each kidney for
  - Obstructed, indeterminate or not obstructed
  - Ranked main rules associated with interpretation
- Tabulated two parameters
  - Were rules identified by expert identified by RENEX?
  - Were other rules identified by RENEX correct?
Justification engine: results

184/203 (91%): RENEX agreed with expert rules
102/103 (99%) correct; secondary importance

Disagreements lead to new knowledge

1. Implement new rules with existing variables
2. New variables need to be considered
3. Adjust certainty factors for rules and variables

Being right for the right reason: Better than just being right?

1. The explanation enhances user confidence because the logical chain is made transparent.
2. Will justification engine make less experienced physicians more experienced?
3. How will physicians actually use the system?
4. Format provides for multilingual implementation.

Prospective study

- 185 kidneys (95 patients)
- *Baseline + furosemide protocol*
- RENEX vs Consensus of 3 experts
## EXPERTS vs RENEX (N = 185)

<table>
<thead>
<tr>
<th>RENEX</th>
<th>EXPERTS</th>
<th>Not Obstructed</th>
<th>Equivocal</th>
<th>Obstructed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not obstructed</td>
<td>101</td>
<td>7</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Equivocal</td>
<td>14</td>
<td>13</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Obstructed</td>
<td>5</td>
<td>9</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>% Correct</td>
<td>84%</td>
<td>45%</td>
<td>92%</td>
<td></td>
</tr>
</tbody>
</table>

AGREEMENT?

How well did RENEX agree with the experts?
How well did the experts agree with each other?
ROC and kappa analysis
ROC curve comparing RENEX to the Consensus [obstructed vs (equivocal + non-obstructed kidneys)]

Certainty factor = +0.2
<table>
<thead>
<tr>
<th>Gold Standard</th>
<th>Expert vs RENEX</th>
<th>Area Under ROC Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Method 1</td>
</tr>
<tr>
<td>Experts 1 &amp; 2</td>
<td>Expert 3 RENEX</td>
<td>0.93</td>
</tr>
<tr>
<td>Experts 1 &amp; 3</td>
<td>Expert 2 RENEX</td>
<td>0.93</td>
</tr>
<tr>
<td>Experts 2 &amp; 3</td>
<td>Expert 1 RENEX</td>
<td>0.92</td>
</tr>
</tbody>
</table>

N = 185
Kappa corrects for chance and quantifies the agreement relative to the maximum possible agreement above chance.

\[
\kappa = \frac{O - E}{1 - E}
\]
Agreement between RENEX and experts (n=185)

<table>
<thead>
<tr>
<th>Experts</th>
<th>Weighted Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert 1 vs Expert 2</td>
<td>0.73</td>
</tr>
<tr>
<td>Expert 1 vs Expert 3</td>
<td>0.65</td>
</tr>
<tr>
<td>Expert 1 vs RENEX</td>
<td>0.71</td>
</tr>
</tbody>
</table>
Limitations of ROC and Kappa

**ROC:**
- Needs gold standard
- Requires two mutually exclusive categories

**Kappa:**
- Summary statistic
- Affected by marginal distribution
- Hard to compare the magnitude (Is 0.72 better than 0.61?)
Mixture distribution analysis
Useful in absence of independent standard
Evaluate agreement in specific classes

Log linear analysis
Direct testing of agreement among raters
RENEX agreed with the experts as well as the experts agreed with each other.

ROC
Kappa analysis
Mixture distribution analysis
Log linear analysis
STATEX
(predictive multivariate model)

- Logistic regression
- Proportional odds modeling
- Interpret diuresis renography scan
- To identify key interpretative variables
- Used Akaike Information Criterion (AIC)
- Baseline + diuretic acquisition
Predictive statistical model

97 patients from archived data base
– Gold standard: consensus of 3 experts

To build the model:
61 patients (120 kidneys)

To test the model:
36 patients (71 kidneys)
<table>
<thead>
<tr>
<th>Baseline variables</th>
<th>Post-furosemide variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAG3 clearance</td>
<td>Time between studies</td>
</tr>
<tr>
<td>Relative uptake</td>
<td>Post-furosemide kidney T1/2</td>
</tr>
<tr>
<td>20 min/2-3 min count ratio</td>
<td>Pelvic T1/2</td>
</tr>
<tr>
<td>20 min/max count ratio</td>
<td>1-2 min/baseline max count ratio</td>
</tr>
<tr>
<td>Baseline kidney Tmax, T ½</td>
<td>20 min/baseline 1-2 min count ratio</td>
</tr>
<tr>
<td>Cortical 20 min/max count ratio</td>
<td>20 min/baseline max count ratio</td>
</tr>
<tr>
<td>Cortical 20 min/2-3 min ratio</td>
<td></td>
</tr>
<tr>
<td>Cortical Tmax, T1/2</td>
<td></td>
</tr>
<tr>
<td>Postvoid/prevoid count ratio</td>
<td></td>
</tr>
<tr>
<td>Postvoid/max count ratio</td>
<td></td>
</tr>
</tbody>
</table>

**Post-furosemide variables**

- Time between studies
- Post-furosemide kidney T1/2
- Pelvic T1/2
- 1-2 min/baseline max count ratio
- 20 min/baseline 1-2 min count ratio
- 20 min/baseline max count ratio
Prediction of normal drainage (no obstruction) on the baseline scan

• 20 min/max count ratio
• Post-void/max count ratio*
• Cortical Tmax
• AUC: 0.93
• Sensitivity 85%, Specificity 93%
Diagnosis of Obstruction

Many variables highly correlated

- Post-void/1-2 min count ratio
- Post-void/max count ratio
- Pelvic T1/2, Whole kidney T1/2
- Post-furosemide
  - 20 min/baseline max count ratio
  - 1-2 min/baseline max count ratio
  - 20 min/baseline 1-2 min count ratio
  - 20 min/baseline max count ratio
Prediction of Obstruction

- Post-furosemide 20 min /baseline max count ratio
- Pelvic T1/2
- AUC I: 0.84, Sens 82%, Spec 78%
- AUC II: 0.84, Sens 83%, Spec 95%
Impact of clinical information on diuretic renography scan interpretation

- 92 patients (baseline + furosemide)
- Camera-based MAG3 clearance
- Consensus of 3 readers
- With and without clinical information
- Scored on continuous scale:
  - Not obstructed: -1.0 to -0.2
  - Equivocal: -0.21 to + 0.19
  - Obstructed: 0.2 to 1.0
Summary and dates

- CT, US, MR
- prior MAG3 scans
- surgeries, stents
- serum creatinine
Impact of clinical information on diuretic renography scan interpretation

- Clinical information changed the category (obstructed, equivocal, non-obstructed) in only 4 of 179 (2.2%) of kidneys.
- The clinical history changed the score >0.1 in only 17 kidneys and changed the score >0.2 in only 3 kidneys.
- *Serum creatinine/MAG3 clearance
Patient: “I would like a second opinion.”
Doctor: “OK. If you need a second opinion, I’ll ask my laptop.”
Add clinical information
Detect and correct for motion
Robust software for ROIs
Software for renal components (pelvis, calyces, parenchyma)
Focal vs diffuse retention
Accommodate other protocols
Renovascular hypertension
Pediatric applications