

Intra- and interobserver variability in interpretation of DMSA scans using a set of standardized criteria

K. Patel, M. Charron, A. Hoberman, M. L. Brown, K. D. Rogers

Division of Nuclear Medicine and Department of Pediatrics, Children's Hospital and University of Pittsburgh Medical Center, Pittsburgh, PA 15213, USA

Received: 24 May 1993/Accepted: 10 August 1993

Abstract. A set of criteria was developed to standardize assessment of DMSA renal scintigraphy which were performed to evaluate children for acute pyelonephritis and renal scarring. This study was undertaken to assess intra- and interobserver variability in the interpretation of DMSA renal scintigraphy using these criteria. Renal contours and parenchyma were assessed in three zones. Contours were assessed as normal or abnormal and parenchymal defects were evaluated in terms of character, shape and degree in three regions (upper and lower pole and midzone). Two nuclear medicine physicians blindly reviewed 57 DMSA scintigraphy on two occasions each. Disagreement of each observer's evaluation of the same scintigraphy on two different occasions was described as intraobserver variability, and the comparison between readings by each of the two observers was described as interobserver variability. High levels of intra- (95.9% and 90.6% respectively, $p < 0.05$) and interobserver agreement (84.4%, $p < 0.05$) were demonstrated. There were minor differences in inconsistencies between the two kidneys or different kidney zones. We conclude that standardization of criteria resulted in higher intra- and interobserver consistency in interpretation of DMSA scintigraphy.

Urinary tract infection (UTI) in young febrile children is an important clinical problem, which when affects the renal parenchyma (acute pyelonephritis) can cause permanent renal damage (scarring). Renal scarring early in life is associated with later development of hypertension and end stage renal disease [1]. Therefore, it is important that UTI be identified promptly and treated appropriately.

Imaging studies have been performed to determine the site of UTI. Radionuclide voiding cystourethrogram, and Tc-99m glucoheptonate, as well as radiographic procedures such as intravenous pyelogram (IVP), ultrasonography (US) and computer tomography (CT) have had

various limitations with respect to sensitivity, accuracy or feasibility [6–8].

More recently, the Tc-99m dimercaptosuccinic acid (DMSA) scintigraphy has been shown to be an accurate method to diagnose acute pyelonephritis in experimental and clinical studies [9–14]. However, various authors have used different criteria for the interpretation of DMSA renal scintigraphy.

In this study, criteria developed by Majd [15], Conway [16], and other authors [17, 18] were unified and modified into a set of standardized criteria for interpreting renal cortical scintigrams (DMSA or glucoheptonate), in an attempt to (1) enhance accuracy and reliability of interpretation, and (2) enable more valid comparison of results of different studies. Intra- and interobserver variability were assessed using these standard criteria in children.

Table 1. Standardized classification for interpretation of DMSA scans

Contours	<ul style="list-style-type: none"> • Normal • Abnormal
Location of parenchymal defects	<ul style="list-style-type: none"> • Upper pole • Mid-zone • Lower pole
Degree of photopenia	<ul style="list-style-type: none"> • Mild • Moderate • Severe • Absent
Shape of defects	<ul style="list-style-type: none"> • Spherical • Wedge • Diffuse
Overall impression	<ul style="list-style-type: none"> • Normal • Scar • Inflammation • Both
Percentage of kidney involvement	<ul style="list-style-type: none"> • 0% • < 10% • 10–24% • 25–49% • 50–74% • > 75%

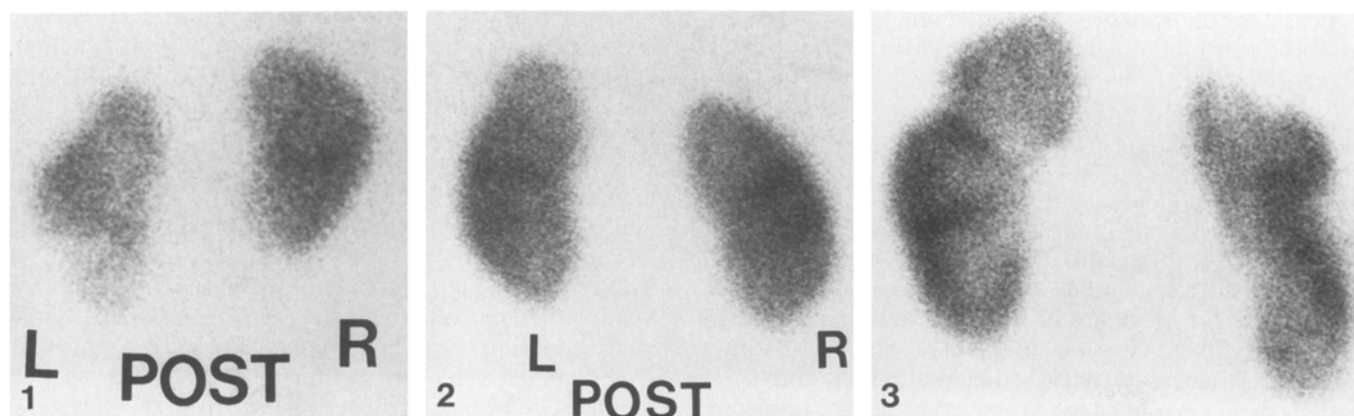


Fig. 1. Acute pyelonephritis in a 6-month-old. There is decreased uptake at the left lower pole with normal contour and no volume loss (inflammation)

Fig. 2. Acute pyelonephritis in a 10-year-old. Note bulging of contour at the left upper pole with homogeneous uptake in the underlying parenchyma (inflammation)

Fig. 3. Diffuse scarring in a 13-year-old. Note areas of decreased uptake with abnormal contour and loss of volume at left midzone and right upper pole and midzone

Methods

Sixty-three consecutive DMSA scans of children suspected of having UTI, vesicoureteral reflux (VUR), or anomalies of the genitourinary system were reviewed. Six studies were excluded on the basis of motion artifacts or poor quality images. Patient age ranged from 14 months to 14 years (mean 8.3 years).

All studies were performed administering 111 MBq/m² (3 mCi/m², minimum 1 mCi and maximum 5 mCi) of Tc-99m DMSA (Mediophysics, Arlington Heights, Illinois). The scintigraphy was acquired 2–3 h after intravenous administration of DMSA, for a maximum of 400 K counts or 5 min, each, in the anterior, posterior and both posterior oblique projections. All images were obtained in analog and digital format, using either a high resolution or an all purpose collimator, on a Siemens Orbitor or Technicare gamma camera. Of the 50 defects noted on the scintigram, 16 were single, 11 were double, and 4 were triple. Nineteen of these involved the upper pole, 15 the midpole, 12 the lower pole, and 3 the medial aspect of the midpole (pelvis).

Two nuclear medicine physicians blindly and independently, reviewed the scintigraphy images on two separate occasions. The period between the two readings ranged from two days to two weeks. The investigators classified their observations according to the criteria listed in Table 1. The area of kidney affected was assessed visually and categorized as percentage of kidney involved. Overall impression was defined as follows:

Normal scintigraphy

- Normal contour, defined as smooth and continuous without indentations
- Homogeneous parenchymal uptake in all regions of both kidneys
- Normal size and reniform shape of both kidneys

Inflammation (acute pyelonephritis) (Fig. 1, 2)

- Slightly bulging or normal contour
- Single or multiple, focal or diffuse areas of decreased activity in

Table 2. Intra- and interobserver agreement in the interpretation of DMSA scans

	Intraobserver observer 1 (%)	Intraobserver observer 2 (%)	Interobserver observers 1 and 2 (%)
Contour	99	97	95
Degree	97	92	86
Character	98	94	89
Shape	98	92	87
Impression	92	85	74
% Renal involvement	90	82	76

the parenchyma, which are diffuse or rarely spherical in shape, in at least 2 projections

- Mild to severe degree of photopenia or rarely complete absence of activity
- No volume loss

Scar (chronic pyelonephritis) (Fig. 3)

- Diffuse or sharp indentation in contour with thinning of cortex
- Any shaped defects with loss of renal volume
- Degree of photopenia more commonly severe or absent activity

Intra- and interobserver agreement were calculated for ratings in the categories listed in Table 1. Weighted kappa analysis was used to determine the statistical significance and strength of associations.

Results

Intraobserver agreement, defined as agreement of each observer's evaluation of the same scintigraphy, blindly, and on two different occasions is reported in Table 2 for observer one and observer two, respectively. The findings of the two readings by each of the two observers were averaged and compared to determine interobserver variability (Table 2). High levels of intra- and interobserver agreement were found for all variables. No differences were statistically significant (for all categories $p < 0.05$). High levels of agreement were observed in the assessment of left kidney compared to right. Similarly, agreement for mid-zone was higher compared to poles. Inconsistencies were more likely to occur in upper poles. However, none of the differences between kidneys or regions were statistically significant. The range of weighted kappa values for observer one (0.8–0.9), for observer two (0.6–

0.8), and for interobserver agreement (0.4–0.7) corroborated the substantial agreement found for the evaluated comparisons.

Discussion

The pediatric kidney is particularly vulnerable to permanent damage secondary to pyelonephritis. Early diagnosis and prompt and adequate treatment of acute infection may reduce the incidence or extent of scarring. An acute phase response consisting of elevated peripheral white blood cell count, erythrocyte sedimentation rate and C-reactive protein, has been used in several studies to indicate infection of the upper urinary tract. However, the studies that evaluated these tests were limited by use of clinical criteria as the validating standard [2, 19]. Until recently, most studies of the significance and long-term consequences of UTI in early childhood have used IVP and/or ultrasonography to identify acute pyelonephritis and renal scarring. The sensitivity of IVP for acute pyelonephritis has been found to be 24–28 % [6]. Serial observations have shown that it can take up to two years for parenchymal scars to become evident by IVP. Even when abnormalities are identified, the extent of renal and perirenal involvement is difficult to determine. In comparison with IVP, DMSA scintigraphy has several additional advantages. Overlying bowel contents can result in poor images with IVP but not with DMSA imaging; DMSA does not cause osmotic overloads or alter renal function; and allergic reactions are rare with DMSA [20–22]. US sensitivity is equal or greater than IVP in diagnosing pyelonephritis [18]. Various studies have reported the sensitivities ranging from 20–40 %. In a study by Bjorgvinsson et al., only 39 % of children with scintigraphically documented pyelonephritis had abnormalities on US. Sensitivity of CT scan is comparable to that of scintigraphy in detecting pyelonephritis; however, its widespread use in children is impractical [7, 10, 23, 24].

No human studies have correlated DMSA scan results with pathological findings. However, Rushton et al. [11], using a piglet model, generated pyelonephritis by surgically induced VUR of infected urine and compared the location and extent of the acute inflammatory response found on histopathological examination with DMSA renal scan findings. DMSA scans had a sensitivity of 87 % and a specificity of 100 % in detecting acute pyelonephritis. In a similar study Arnold et al., using a similar model, observed a sensitivity of 85 % and a specificity of 97 % in detecting macroscopic renal scarring [12]. Jakobsson et al. reported a sensitivity of 92 % for DMSA scans in children during acute UTI [14].

In a recent review article [25], Andrich and Majd stated that infants and children who have a UTI accompanied by fever and toxicity cannot properly be diagnosed as having acute pyelonephritis on the basis of clinical signs and symptoms or laboratory parameters alone. They considered DMSA scintigraphy to be the “gold standard” for identification of the renal parenchymal changes of acute pyelonephritis and recommended its use as the primary study for diagnosis. Additional attractions of DMSA renal

scans are that they (1) are readily available and relatively inexpensive, (2) allow quantitation of renal function, (3) are non-invasive, (4) yield high resolution images, (5) can detect obstruction when combined with ^{99m}Tc Mag-3 scintigraphy.

Various investigators have used different criteria for interpretation of renal cortical scintigraphy (DMSA or glucoheptonate). Majd et al. described the acute pyelonephritic lesion as an area of diminished cortical uptake with or without a bulging contour, not associated with volume loss, or the less common diffuse decreased uptake in an enlarged kidney [15]. Handmaker described a relatively specific pattern in acute pyelonephritis as a flare shaped region of decreased activity radiating from the pelvicalyceal system towards the periphery of the kidney [17]. Sty et al. defined spherical regions of decrease activity as inflammatory lesions, and also noted them to be frequently located in the poles [18]. In different studies by Majd and Conway, a scar was described as loss of activity with indentation of contour and loss of volume of the involved cortex. It has also been noted that defects that initially are extensive and extend to the cortical surface are more prone to progress to scars [15, 16].

The criteria for classifying DMSA scans developed and used in this study were successful in standardizing interpretation of scans with respect to location, impression (pyelonephritis and/or scar), and severity of disease. Both intra- and interobserver agreement in the interpretation of DMSA scans were high. Degree of photopenia and percentage of renal involvement had four and six categories, respectively. This narrow separation between categories may explain some inconsistencies. In most cases, when there was observer disagreement, it occurs between categories in immediately adjacent groups. For example, 24 defects were described as slight photopenia on one observation and normal on the second reading; 10 defects were described as moderate photopenia on the first reading and severe photopenia on the second. Similarly, in percentage of renal involvement, 5 kidneys had less than 10 % involvement on the initial evaluation, and 10–25 % involvement on the second one. There was excellent agreement regarding character and shape with only coincidental variations. The main cause of variability in overall impression was the description as both inflammation and scar on the initial reading and either scar or inflammation on the subsequent reading.

In the interobserver comparisons, 35 inconsistencies in percentage of renal involvement also were between adjacent categories. Accordingly, a change of at least two categories might be better to define a change between two readings. The most frequent inconsistency in overall impression was that one observer diagnosed small contracted kidneys as hypoplastic, while the other observer described them as being scarred. Perhaps this constitutes a limitation of the criteria, because a clear description of a hypoplastic kidney is often difficult. However, scan findings together with clinical parameters helped to differentiate hypoplastic from scarred kidneys. Another finding was that one observer had a greater tendency to define defects as both scar and inflammation, whereas the other observer classified them as either scar or inflammation. A

possible explanation for this discrepancy in both intra- and interobserver variability, may be that some defects appear as scars in one projection and inflammation in another. Therefore, evaluating multiple projections before describing a defect is important. Additionally, defects in the medial aspects of three kidneys were called inflammation by one observer, while the other observer described them as normal collecting systems. In both intra- and interobserver comparisons of location, greater inconsistency in interpretations of the upper poles, when it occurred, may have been attributable to liver attenuation on the right side, to splenic attenuation on the left in the anterior projection and differences in kidney depth of the upper versus lower poles.

The findings of this study suggest that high intra- and interobserver consistency can be achieved in interpreting DMSA scans when using a standardized set of criteria. These criteria, modified to reduce inconsistencies, currently are being used in a study of treatment of pyelonephritis, to evaluate DMSA scan images obtained with pinhole collimation known to provide the highest possible resolution.

References

1. Winberg J (1987) Clinical aspects of urinary tract infection. Chapter 36. In: Holliday MA, Banat TM, Vernier RL (eds) *Pediatric nephrology*. Williams and Williams, Baltimore
2. Marild S, Hellstrom M, Jodal U, Svanborg Eden C (1989) Fever, bacteriuria and concomitant disease in children with urinary tract infection. *Pediatr Infect Dis J* 8: 36
3. Fairley KF, Carson NE, Gutch RC et al. (1971) Site of infection in acute urinary tract infection in general practice. *Lancet* II: 615
4. Eykyn S, Lloyd-Davies RW, Shuttleworth KED, Vinnicombe J (1972) The localization of urinary tract infection by ureteric catheterization. *Invest Urol* 9: 271
5. Whitaker J, Hewstone AS (1969) The bacteriologic differentiation between upper and lower urinary tract infection in children. *J Pediatr* 74: 364
6. Harrison RB, Shaffer HA (1979) The roentgenographic findings in acute pyelonephritis. *JAMA* 241: 1718
7. June CH, Browning MD, Smith CP et al. (1985) Ultrasonography and computed tomography in severe urinary tract infection. *Arch Intern Med* 145: 841-845
8. Traisman ES, Conway JJ, Traisman HS et al (1986) The localization of urinary tract infection with ^{99m}Tc glucoheptonate scintigraphy. *Pediatr Radiol* 16: 403
9. Verber IG, Strudley MR, Meller ST (1988) $^{99\text{Tc}}$ dimercaptosuccinic acid (DMSA) scan as first investigation of urinary tract infection. *Arch Dis Child* 63: 1320
10. Bjorgvinsson E, Majd M, Eggli KD (1991) Diagnosis of acute pyelonephritis in children: comparison of sonography and $^{99\text{mTc}}$ DMSA scintigraphy. *Am J Roentgenol* 157: 539
11. Rushton HG, Majd M, Vim D (1988) Evaluation of $^{99\text{Tc}}$ Technetium-dimercaptosuccinic acid renal scans in experimental acute pyelonephritis in piglets. *J Urol* 140: 1169
12. Arnold AJ, Brownless SM, Carty HM, Rickwood AMK (1990) Detection of renal scarring by DMSA scanning - an experimental study. *J Pediatr Surg* 25: 391
13. Rushton HG, Majd M, Jantusch B, Wiedermann BL, Belman AB (1992) Renal scarring following reflux and nonreflux pyelonephritis in children: evaluation with $^{99\text{Tc}}$ technetium-dimercaptosuccinic acid scintigraphy. *J Urol* 147: 1327
14. Jakobsson B, Nolstedt L, Svensson L, Soderlundh S, Berg U (1992) $^{99\text{mTc}}$ Technetium-dimercaptosuccinic acid scan in the diagnosis of acute pyelonephritis in children: relation to clinical and radiological findings. *Pediatr Nephrol* 6: 328
15. Majd M, Rushton HG (1992) Renal cortical scintigraphy in the diagnosis of acute pyelonephritis. *Semin Nucl Med* 22: 98
16. Conway JJ (1988) The role of scintigraphy in urinary tract infection. *Semin Nucl Med* 18: 308
17. Handmaker H (1982) Nuclear renal imaging in acute pyelonephritis. *Semin Nucl Med* 12: 246
18. Sty JR, Wells RG, Starshak RJ, Schroeder BA (1980) Imaging in acute renal infection in children. *Am J Roentgenol* 148: 471
19. Pylkkanen J, Vilska J, Koskimies O (1981) The value of diagnosis of childhood urinary tract infection in predicting renal injury. *Acta Paediatr Scand* 70: 879
20. Goldraich NP, Ramos OL, Goldraich IH (1989) Urography versus DMSA scan in children with vesicoureteric reflux. *Pediatr Nephrol* 3: 1
21. Little PJ, McPherson DR, DeWardener HE (1965) The appearance of the intravenous pyelogram during and after pyelonephritis. *Lancet* I: 1186
22. Shaw P, Whitear P, Gordon I (1987) Comparison of DMSA scans and IVU's. *Pediatr Radiol* 17: 341 (abstr)
23. Dinkel E, Orth S, Dittrich M, Schulte-Wisserman H (1986) Renal sonography in the differentiation of upper from lower urinary tract infection. *AJR* 146: 775
24. Kahn PC (1979) Renal imaging with radionuclides, ultrasound and computed tomography. *Semin Nucl Med* 9: 43
25. Andrich MP, Majd M (1992) Diagnostic imaging in the evaluation of the first urinary tract infection in infants and young children. *Pediatrics* 90: 436