ASSESSING THE EVIDENCE FOR DIAGNOSTIC IMAGING:
GOING BEYOND ACCURACY

Martin H Reed, MD FRCPC
Chair, Guidelines Working Group
Canadian Association of Radiologists
Head, Department of Diagnostic Imaging
Children’s Hospital, Winnipeg, Canada
GOALS

1. Levels of evidence in diagnostic imaging
2. Issues with levels of evidence
3. Accuracy in diagnostic imaging
4. Problems with accuracy
5. Going beyond accuracy
A hierarchical model of efficacy

- Level 1: Technical efficacy
- Level 2: Diagnostic accuracy efficacy
- Level 3: Diagnostic thinking efficacy
- Level 4: Therapeutic efficacy
- Level 5: Patient outcome efficacy
- Level 6: Societal efficacy
A hierarchical model of efficacy

• Level 1: Technical efficacy
  Level 1. Technical efficacy
  Resolution of line pairs
  Modulation transfer function change
  Gray-scale range
  Amount of mottle
  Sharpness

• Level 2: Diagnostic accuracy efficacy

• Level 3: Diagnostic thinking efficacy

• Level 4: Therapeutic efficacy

• Level 5: Patient outcome efficacy

• Level 6: Societal efficacy
A hierarchical model of efficacy

- Level 1: Technical efficacy
- Level 2: Diagnostic accuracy efficacy
- Level 3: Diagnostic thinking efficacy
- Level 4: Therapeutic efficacy
- Level 5: Patient outcome efficacy
- Level 6: Societal efficacy
A hierarchical model of efficacy

Level 5. Patient outcome efficacy
  Percentage of patients improved with test compared with without test
  Morbidity (or procedures) avoided after having image information
  Change in quality-adjusted life expectancy
  Expected value of test information in quality-adjusted life years (QALYs)
  Cost per QALY saved with image information

Level 6. Societal efficacy
  Benefit–cost analysis from societal viewpoint
  Cost–effectiveness analysis from societal viewpoint
A hierarchical model of efficacy

- Level 1: Technical efficacy
- Level 2: Diagnostic accuracy efficacy
  - Yield of abnormal or normal diagnoses in a case series
  - Diagnostic accuracy (percentage correct diagnoses in case series)
  - Predictive value of positive or negative examination (in a case series)
  - Sensitivity and specificity in a defined clinical problem setting
- Level 3: Diagnostic thinking efficacy
- Level 4: Therapeutic efficacy
- Level 5: Patient outcome efficacy
- Level 6: Societal efficacy
Revised estimates of diagnostic test sensitivity and specificity in suspected biliary tract disease.


Department of Medicine, University of Pennsylvania, Philadelphia.

Comment in:
ACP J Club. 1995 May-Jun;122(3):76.

Abstract

BACKGROUND: The purpose of this study was to estimate the sensitivity and specificity of diagnostic tests for gallstones and acute cholecystitis.

METHODS: All English-language articles published from 1966 through 1992 about tests used in the diagnosis of biliary tract disease were identified through MEDLINE. From 1614 titles, 668 abstracts were examined and 322 articles were read to identify 61 articles with information about sensitivity and specificity. Application of exclusion criteria based on clinical and methodologic criteria left 30 articles for analysis. Cluster-sampling methods were adapted to obtain combined estimates of sensitivities and specificities. Adjustments were made to estimates that were biased because the gold standard was applied preferentially to patients with positive test results.

RESULTS: Ultrasound has the best unadjusted sensitivity (0.97, 95% confidence interval, 0.95 to 0.99) and specificity (0.96, 95% confidence interval, 0.95 to 1.00) for evaluating patients with suspected gallstones. Adjusted values are 0.84 (0.76 to 0.92) and 0.99 (0.97 to 1.00), respectively. For evaluating patients with acute cholecystitis, ultrasound has the best sensitivity (0.97, 95% confidence interval, 0.96 to 0.98) and specificity (0.90, 95% confidence interval, 0.86 to 0.95) for evaluating patients with suspected acute cholecystitis. Test performance is unaffected by delayed imaging. Unadjusted sensitivity and specificity of ultrasound in evaluating patients with suspected acute cholecystitis are 0.94 (0.92 to 0.96) and 0.78 (0.01 to 0.96); adjusted values are 0.88 (0.74 to 1.00) and 0.80 (0.62 to 0.98).

CONCLUSIONS: Ultrasound is superior to oral cholecystogram for diagnosing cholelithiasis, and radionuclide scanning is the test of choice for acute cholecystitis. However, sensitivities and specificities are somewhat lower than commonly reported. We recommend estimates that are midway between the adjusted and unadjusted values.
GUIDELINE

? Gallstones
Do ultrasound
FRACTURES
Level 2: ACCURACY
FRACTURES
Level 2: ACCURACY

What is the accuracy of an x-ray for the diagnosis of a fracture?
FRACTURES
Level 2: ACCURACY

Diagnosis of acute fractures of the extremities: comparison of low-field MRI and conventional radiography.
Remplik P, Stabler A, Merl T, Rehner F, Behndorf K
Department of Radiology, Klinikum Augsburg, Stenglinstrasse 2, 86156 Augsburg, Germany. ph.remplik@gmx.de

Abstract
The aim of this study was to compare low-field MRI (0.2 T) and conventional radiography for the detection of acute fractures of the distal part of the extremities. X-ray and MRI examinations of 78 (41 fractures, 37 without fracture) patients with the clinical suspicion of an acute fracture in the distal part of the extremities were compared. Four experienced radiologists, two for each of the two modalities, independently analyzed the images. Interobserver variability and receiver operating characteristic (ROC) analysis for both methods were established. The MRI and conventional radiography revealed an accuracy of 81.4% and of 79.5%, respectively, in the detection of acute fractures. The diagnostic accuracy of MRI to detect fractures in the hand and foot proved to be significantly inferior to conventional X-ray examinations. On the other hand, MRI achieved a better accuracy for the examination of bones near a large joint. The interobserver variability for both methods was rated as moderate. In ROC analysis both methods were rated as good. There was no statistical difference of the accuracy between low-field MRI and conventional radiography in the detection of acute fractures of the distal part of the extremities. Consequently, a routine use of low-field MRI as an alternative to conventional radiography to diagnose acute fractures of the extremities seems not to be justified.
PNEUMONIA
Level 2: ACCURACY
What is the accuracy of an x-ray for the diagnosis of pneumonia?
Conclusions: We have found that studies assessing the diagnostic accuracy of clinical, radiological, and laboratory tests for bacterial childhood pneumonia have used a heterogeneous group of gold standards, and found, at least in part because of this, that index tests have widely different accuracies. These findings highlight the need for identifying a widely accepted gold standard for diagnosis of bacterial pneumonia in children.
PROBLEMS WITH ACCURACY

1. Not known
2. Not enough
A hierarchical model of efficacy

• Level 1: Technical efficacy
• Level 2: Diagnostic accuracy efficacy
• Level 3: Diagnostic thinking efficacy
• Level 4: Therapeutic efficacy
• Level 5: Patient outcome efficacy
• Level 6: Societal efficacy
FRACTURES
Level 2: ACCURACY
Computed tomography in head trauma.

Dublin AB, French BN, Rennick JM.

Abstract

Retrospective analysis of 200 cases of documented head trauma demonstrated an accuracy approaching 100% in the diagnosis of intra- and extracerebral collections of blood. Caution must be exercised in the evaluation of trauma 1 to 5 weeks old, since subdural hematomas have the same density as normal brain tissue, and angiography may be necessary. The clinical diagnosis of brainstem contusion is associated with a remarkably high level (54%) of surgically correctable lesions. The use of computed tomography in the evaluation of other traumatic intracranial lesions is discussed.
HEAD TRAUMA PROBLEMS WITH ACCURACY

Variation in utilization of computed tomography scanning for the investigation of minor head trauma in children: a Canadian experience.


RESULTS: One thousand one hundred sixty-four children were included in the study. One hundred seventy-one (15%) had a CT scan, of which 60 (35%) were abnormal. There was a significant difference in the rate of ordering of CT scans among the participating hospitals, but no significant difference in the rate of abnormal CT scans. Mechanism of injury, GCS, and loss of consciousness were significantly related to the presence of an abnormal CT scan.

Figure 1. Case severity-adjusted rate of order computed tomography (CT) scan by hospital.
HEAD TRAUMA

Minor head trauma in a child
To CT or not to CT
Level 3: DIAGNOSTIC THINKING EFFICACY GUIDELINE (DECISION RULE)

CATCH: a clinical decision rule for the use of computed tomography in children with minor head injury.


RESULTS: Among the 3868 patients enrolled (mean age 9.2 years), 95 (2.5%) had a score of 13 on the Glasgow Coma Scale, 282 (7.3%) had a score of 14, and 3469 (90.2%) had a score of 15. CT revealed that 159 (4.1%) had a brain injury, and 24 (0.6%) underwent neurologic intervention. We derived a decision rule for CT of the head consisting of four high-risk factors (failure to reach score of 15 on the Glasgow coma scale within two hours, suspicion of open skull fracture, worsening headache and irritability) and three additional medium-risk factors (large, boggy hematoma of the scalp, signs of basal skull fracture, dangerous mechanism of injury). The high-risk factors were 100.0% sensitive (95% CI 86.2%-100.0%) for predicting the need for neurologic intervention and would require that 30.2% of patients undergo CT. The medium-risk factors resulted in 98.1% sensitivity (95% CI 94.6%-99.4%) for the prediction of brain injury by CT and would require that 52.0% of patients undergo CT.
APPENDICITIS
Level 2: ACCURACY


Unenhanced limited CT of the abdomen in the diagnosis of appendicitis in children: comparison with sonography.

Lawe LH, Penney MW, Stein SM, Heller RM, Neblett WW, Shyr Y, Homanz-Schulman M.

Department of Radiology and Radiological Sciences, Vanderbilt University Children's Hospital and Medical Center, D-1120 Medical Center North, Nashville, TN 37232-2675, USA.

Abstract

OBJECTIVE: The purpose of this investigation is to determine the sensitivity, specificity, and accuracy of unenhanced limited CT of the abdomen in children with suspected appendicitis and compare these results with graded compression sonography.

MATERIALS AND METHODS: Seventy-six children underwent unenhanced limited CT over a 11-month period for evaluation of suspected appendicitis. A historical cohort of 88 consecutive children who had undergone graded compression sonography was identified. Results were correlated with surgical, pathologic, chart, and clinical follow-up data. The sensitivity, specificity, accuracy, rate of alternate diagnosis, time to perform examinations, and charge at our institution were determined for unenhanced limited CT and sonography.

RESULTS: Sensitivity, specificity, and accuracy for unenhanced limited CT were 97%, 100%, 99%, respectively, and were 100%, 88%, 91%, respectively, for sonography. Alternate diagnoses were suggested in 35% and 23%
APPENDICITIS

CT
or
Ultrasound
Unenhanced limited CT of the abdomen in the diagnosis of appendicitis in children: comparison with sonography.

Lowe LH, Penney MW, Stein SM, Heller RM, Nebbett WW, Shyr Y, Herranz-Schulman M

Department of Radiology and Radiological Sciences, Vanderbilt University Children’s Hospital and Medical Center, D-1120 Medical Center North, Nashville, TN 37232-2675, USA.

Abstract

OBJECTIVE: The purpose of this investigation is to determine the sensitivity, specificity, and accuracy of unenhanced limited CT of the abdomen in children with suspected appendicitis and compare these results with graded compression sonography.

MATERIALS AND METHODS: Seventy-six children underwent unenhanced limited CT over a 11-month period for evaluation of suspected appendicitis. A historical cohort of 88 consecutive children who had undergone graded compression sonography was identified. Results were correlated with surgical pathology, chart, and clinical follow-up data. The sensitivity, specificity, accuracy, rate of alternate diagnosis, time to perform examinations, and charge at our institution were determined for unenhanced limited CT and sonography.

RESULTS: Sensitivity, specificity, and accuracy for unenhanced limited CT were 97%, 100%, 99%, respectively, and were 100%, 88%, 91%, respectively, for sonography. Alternate diagnoses were suggested in 36% and 26% children without appendicitis who had unenhanced limited CT and sonography, respectively. Unenhanced limited CT required 5 min and sonography required 20-30 min to perform. The charge at our institution was $408 for unenhanced limited CT and $285 for sonography.

CONCLUSIONS: CT can be performed rapidly in children without IV, oral, or rectal contrast medium. Unenhanced limited CT and sonography are highly sensitive, specific, and accurate in the evaluation of children with suspected appendicitis.
APPENDICITIS
GUIDELINE - Values

Cost
vs.
Time
vs.
Radiation
APPENDICITIS GUIDELINE

?Appendicitis in a child
Do ultrasound first
Bronchiolitis of infancy is a clinically diagnosed respiratory condition presenting with breathing difficulties, cough, poor feeding, irritability and, in the very young, apnoea. These clinical features, together with wheeze and/or crepitations on auscultation combine to make the diagnosis. Bronchiolitis most commonly presents in infants aged three to six months.¹

Bronchiolitis occurs in association with viral infections (respiratory syncytial virus; RSV, in around 75% of cases)² and is seasonal, with peak prevalence in the winter months (November to March) when such viruses are widespread in the community. Re-infection during a single season is possible.
BRONCHIOLITIS

An infant with bronchiolitis
Do I order a chest x-ray?
Diagnosis and Management of Bronchiolitis

Subcommittee on Diagnosis and Management of Bronchiolitis

Endorsed by the American Academy of Family Physicians, the American College of Chest Physicians, and the American Thoracic Society.

RECOMMENDATION 1a
Clinicians should diagnose bronchiolitis and assess disease severity on the basis of history and physical examination. Clinicians should not routinely order laboratory and radiologic studies for diagnosis (recommendation: evidence level B; diagnostic studies with minor limitations and observational studies with consistent findings; preponderance of benefits over harms and cost).
A hierarchical model of efficacy


• Level 1: Technical efficacy
• Level 2: Diagnostic accuracy efficacy
• Level 3: Diagnostic thinking efficacy
• Level 4: Therapeutic efficacy
• Level 5: Patient outcome efficacy
• Level 6: Societal efficacy
PROBLEMS WITH ACCURACY

1. Not known
2. Not enough
A hierarchical model of efficacy

• Level 1: Technical efficacy
• Level 2: Diagnostic accuracy efficacy
• Level 3: Diagnostic thinking efficacy
• Level 4: Therapeutic efficacy
• Level 5: Patient outcome efficacy
• Level 6: Societal efficacy