Original article

Radiographer-acquired and radiologist-reviewed ultrasound examination – agreement with radiologist's bedside evaluation

Carina Stenman[1,2], Lars Thorelius[3], Anders Knutsson[2] and Örjan Smedby[1,4]

[1]Division of Radiological Sciences, Radiology, IMH, Linköping University, Linköping, Sweden; [2]Radiology, Linköping University Hospital, Linköping, Sweden; [3]Department of Radiology, Copenhagen University Hospital at Herlev, Denmark; [4]Center for Medical Image Science and Visualization (CMIV), Linköping University, Linköping, Sweden
Correspondence to: Carina Stenman. Email: carina.stenman@liu.se

Abstract

Background: Growing demand for ultrasound examinations and higher quality requirements motivate searching for routines combining the diagnostic accuracy of radiologist-performed examinations with the economical advantages of sonographer-performed examinations. One possible approach is to use strictly standardized acquisition and documentation schemes that give the radiologist access to all relevant information after the examination.

Purpose: To compare a recently introduced routine, combining acquisition by a radiographer, documentation as standardized cine-loops, and review by a radiologist ('standardized method'), with the formerly used routine where the diagnosis is made bedside by the radiologist ('traditional method').

Material and Methods: In 64 policlinic patients, the kidneys ($n = 27$) or the gallbladder ($n = 37$) were examined with both the standardized and the traditional method. The radiologists' findings of hydronephrosis, tumors, cysts, echogenicity changes, and cortical thickness (in the kidneys), and wall thickness, concrements, and polyps (in the gallbladder) were compared between the methods with respect to agreement (proportion of agreement and kappa coefficient) as well as systematic differences (McNemar's test).

Results: The findings at the gallbladder examination showed a median agreement of 97% (86–100%; kappa = 0.64–1.00), and those at the kidney examination, an agreement of 90% (78–100%; kappa = 0.69–1.00). There were no significant systematic differences between the methods.

Conclusion: The satisfactory agreement in this preliminary study indicates that the new workflow with ultrasound examinations performed by a radiographer and analyzed off-line by a radiologist is promising, and motivates further studies.

Keywords: Ultrasound, radiographer, radiologist, comparison, diagnostic

Submitted May 20, 2009; accepted for publication June 20, 2010

Ultrasound examinations have become increasingly common for investigating the abdomen. This may depend on the technical development of ultrasound equipment, resulting in improved image quality, as well as advantages such as lack of radiation, availability, safety, low cost, and non-invasiveness. As the great demand for ultrasound examinations can lead to capacity problems for radiological services, an extended role of the radiographer has been discussed (1, 2).

The department of Radiology at the University hospital in Linköping, Sweden, uses a standardized method for ultrasound examinations. The examination is performed by a radiologist or, in suitable cases, by a radiographer, according to an organ scheme and then stored as cine-loops. The standardized method facilitates comparisons between old and new examinations of the same patient. The dynamic scans are saved in the Picture Archiving and Communication System (PACS) from where the films can be retrieved and reviewed on a later occasion.

A prerequisite for recommending this way of working for general use is that no diagnostic information is lost in the process. Therefore, the aim of this study was to evaluate the agreement between the proposed method and the method traditionally used in many radiology departments, where radiologists perform the examination.

Material and Methods

Sixty-four adult consecutive outpatients referred for clinical abdominal ultrasound examination at the Radiology department of the University hospital in Linköping, Sweden, were included between October and December 2006. In the examinations focusing on the gallbladder (in accordance with the clinical question), the left lobe of the liver was also included. Each patient was examined by one radiographer and one radiologist during the same visit. First, a radiographer performed the examination according to the standardized method, and then a radiologist performed an examination using the traditional method. The two radiologists participating in the study had 12–17 years' experience of radiological ultrasound and were both familiar with the two examination methods. The radiographer had worked with ultrasound for about two years and had only used the standardized method. All sonograms were obtained with ACUSON Sequoia (Siemens Medical Systems, Erlangen, Germany), using 6C2 and 4C1 convex transducers. The technical parameters were optimized on a case-by-case basis.

The cine-loops used with the standardized method were dynamic clips, with the scan covering 5–10 cm in 5–10 seconds. The dynamic documentation included both longitudinal and transversal views covering the whole organ or region of interest. Documentation was always made from cranial to caudal and from left to right, independent of patient position. The examination followed an organ scheme, including about 10 sweeps, depending on the organ to be examined. Thus, all patients were examined in a similar fashion (3). Static images were used only for measurement.

The radiographer examined the patient using the standardized method, and stored the dynamic films in the PACS (Syngo Dynamics, Siemens Medical Systems, Erlangen, Germany). One of the radiologists then examined the same patient using the traditional method. The radiologist who examined the patient with the traditional method was never the same as the one who reviewed the standardized examination made by the radiographer. All standardized examinations were reviewed at a workstation, where the speed of the cine-clips was 17 images per second. A body marker and a transducer marker on the screen were used for informing the radiologist about the positioning of the patient and of the transducer relative to the patient.

Standardized forms describing the findings were filled out by the radiologists immediately after examining the patient with the traditional method and when reviewing the standardized examination made by the radiographer. In the kidneys, the findings included cortical thickness (normal or decreased), hydronephrosis (presence or absence of dilated calyceal system as judged visually), echogenicity (normal or increased; judged visually) and tumor (presence or absence), as well as the number of simple cysts (0; 1–2; 3–7 or >7) and their maximum diameter ($\leq$2 cm; 2–5 cm or >5 cm). In the gallbladder, thickness of gallbladder wall and the number and size ($\leq$5 mm or >5 mm) of concrements and polyps were recorded.

The findings obtained with both methods were then compared. The agreement between the two methods was calculated in percent and as the kappa statistic (4). The 95% confidence limits for the agreement were calculated with exact computation from the binomial distribution, and those for kappa with the customary normal approximation. In cases where the observed number of positive or negative findings with either method was less than 3, kappa or confidence limits for the agreement were not calculated (5). For dichotomous variables, McNemar's test with exact computation from the binomial distribution was used to determine whether there was a systematic difference resulting in a higher frequency of positive findings with either of the methods. The numbers and sizes of cysts, concrements and polyps were compared between the methods with the kappa statistics and agreement.

Results

Of the 64 patients, 30 were men and 34 women. The age range was 19–93 years (median 60 years). In 27 cases, the kidneys were examined, and in 37, the gallbladder.

Examples of frames from the cine-loops of the standardized method are seen in Figs. 1 and 2. The frequency of pathological findings in the kidneys with the traditional and the standardized method is shown in Table 1. The most common finding was simple renal cysts. For none of the findings was there a significant difference in frequency between the two methods. The agreement between the two methods varied between 78–100%, and corresponding kappa values between 0.69–1.00 (Table 2). The lowest agreement was found for increased echogenicity of the renal parenchyma and for the size of cysts.

Findings not described in the predetermined protocol were seen in six of the 27 patients whose kidneys were examined. In two cases, the findings agreed between the two methods: a large prostate gland and liver metastases. In three cases, there were findings that were seen only when reviewing the standardized method: a bladder tumor and a small concrement in the right kidney in one patient, and thinning of the parenchyma in the left and the right kidney in another patient. In one case, enlargement of the prostate gland was noticed only with the traditional method.

In 22 cases, the length of the kidneys recorded with the standardized method was 0.5–1.0 cm smaller than with the traditional method. In three cases there was exact agreement. Two examinations showed a slightly greater kidney length (0.5–1.0 cm) with the standardized method.

Frequencies of pathological findings in the gallbladder are given in Table 3. Again, no significant differences in frequency between the methods were found. For these findings, the agreement varied between 86–100%, and kappa between 0.64–1.00, with the lowest values for the number of concrements and size of polyps, respectively (Table 4).

In one case a finding of sludge in the gallbladder was made only with the standardized method.

Discussion

In this paper we describe a novel way of evaluating ultrasound examinations, using stored standardized cine-loops

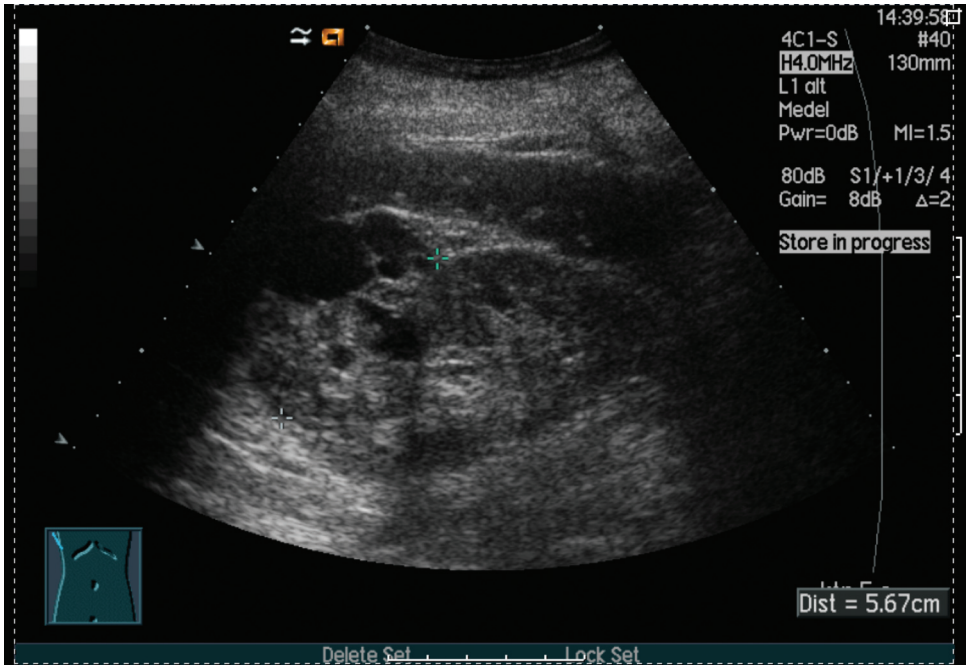

Fig. 1 Mixed tumor in kidney. Single oblique coronal frame from a sweep in the longitudinal direction

that let the radiologist make the evaluation after the examination. This approach, in contrast to documentation with static images, can also be used to let radiographers perform the examination while the diagnostic work is still the task of the radiologist. However, before the new approach is introduced on a large scale, its diagnostic performance needs to be evaluated.

The main aim of this study was to compare the newly introduced standardized method with that traditionally used in most Swedish radiology departments, i.e. bedside examination by a radiologist. The findings indicate good agreement in particular for gallbladder examinations, but also for kidney examinations. All renal findings had at least 78% agreement and kappa of at least 0.69, which seems satisfactory. However, there is considerable uncertainty in our estimates of agreement and kappa, as illustrated by wide confidence limits, due to the small sample size of this study.

Using McNemar's test for matched data, we found no significant systematic differences between the two methods, i.e.

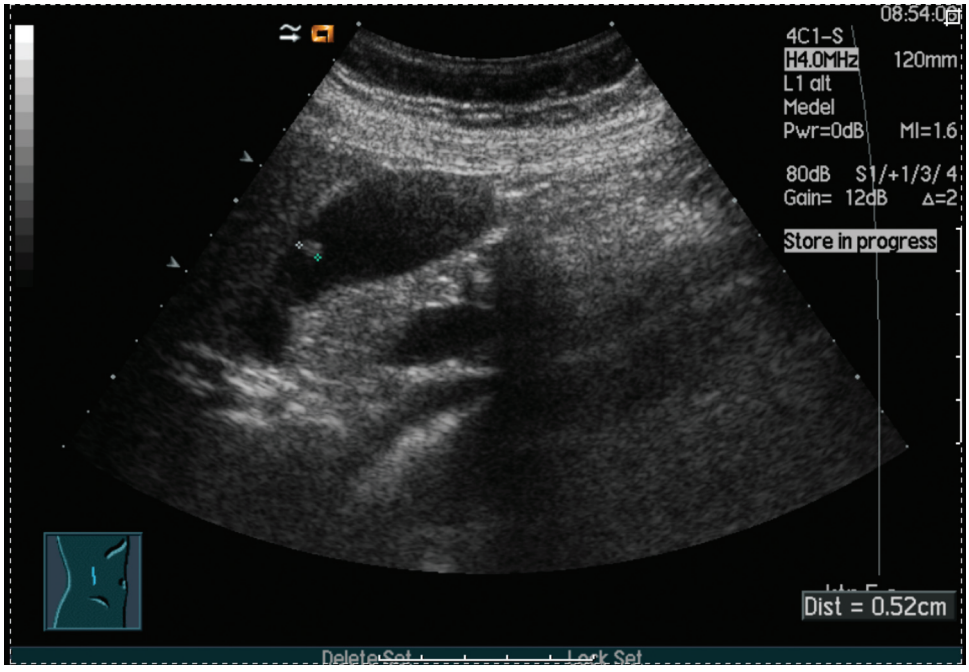

Fig. 2 A polyp in the gallbladder with a maximum diameter of 0.5 cm. Single coronal frame from a sweep in the longitudinal direction

Table 1. Frequency of pathological findings in the kidneys with traditional and standardized method in 27 patients

Finding	Right kidney		Left kidney	
	Patients with positive finding with traditional method (*n*)	Patients with positive finding with standardized method (*n*)	Patients with positive finding with traditional method (*n*)	Patients with positive finding with standardized method (*n*)
Decreased cortical thickness	5	6	3	1
Hydronephrosis	4	3	2	3
Increased echogenicity	2	4	1	5
Tumor	1	1	1	1
Cysts	11	11	10	11
Cysts >2 cm	5	4	4	2

Table 2. Agreement between traditional and standardized method and corresponding kappa values for findings in the kidneys

Finding	Right kidney		Left kidney	
	Agreement, % (95% confidence limits)*	Kappa (95% confidence limits)*	Agreement, % (95% confidence limits)*	Kappa (95% confidence limits)*
Decreased cortical thickness	89 (71%; 97%)	0.69 (0.12; 1.00)	92	–
Hydronephrosis	96 (81%; 99%)	0.83 (0.14; 1.00)	96	–
Increased echogenicity	78	–	78	–
Tumor	100	–	100	–
Presence of cysts	85 (63%; 95%)	0.70 (0.23; 1.00)	89 (70%; 97%)	0.76 (0.12; 1.00)

*Confidence limits and Kappa values were not calculated when observed numbers were less than 3

none of the findings was found to occur significantly more often with one of the methods. However, it cannot be excluded that such systematic differences may be revealed by future studies with a larger number of patients. As for the length of the kidneys, the standardized method in general yielded values 0.5–1.0 cm smaller than the traditional method. This can be explained by the fact that in the traditional method, the angle of the transducer is chosen in a way optimal for finding the maximum length of the kidney, whereas the standardized method uses planes relative to the patient's orientation.

New findings not described in the predetermined protocol of the kidneys included three findings only seen when reviewing the standardized exam. This may be related to the fact that the radiologist had the opportunity to review the exam several times. Most important is that the findings of serious pathology are not missed by either method. In the kidneys, there was 100% agreement for tumor findings (Fig. 1), but it remains to be seen whether this result can be reproduced in a larger material.

The kappa coefficient is commonly used for measuring agreement in radiological studies (6). In some cases in Tables 2 and 4, however, there was a discrepancy between rather high agreement figures and low kappa values. This may occur if there is an imbalance between positive and negative findings in the 2×2 tables. It has, therefore, been recommended to supplement the kappa values with values of percent agreement (5, 6).

The increasing demand for ultrasound examinations can lead to long waiting times (2). The possibility to let radiographers rather than radiologists perform the examination may help to solve this problem. In our clinical practice we have observed that a radiologist can evaluate around 10 examinations per hour that have been performed by a

Table 3. Frequency of pathological findings in the gallbladder with traditional and standardized method in 37 patients

Finding	Patients with positive finding with traditional method (*n*)	Patients with positive finding with standardized method (*n*)
Increased thickness of gallbladder wall	2	1
Concrements in gallbladder	12	12
Concrements in gallbladder >5 mm	8	8
Polyps in gallbladder	3	3
Polyps in gallbladder >5 mm	0	0

Table 4. Agreement between traditional and standardized method and corresponding kappa values for findings in the gallbladder

	Agreement, % (95% confidence limits)*	Kappa (95% confidence limits)*
Increased thickness of gallbladder wall	97	–
Presence of concrements	100 (75%; 100%)	1.00 (0.63; 1.00)
Number of concrements	86 (82%; 99%)	0.82 (0.44; 1.00)
Size of concrements	100 (90%; 100%)	1.00 (0.64; 1.00)
Presence of polyps	94 (81%; 99%)	0.64

*Confidence limits and Kappa values were not calculated when observed numbers were less than 3

radiographer, whereas performing the examination typically take 30 minutes of radiologist time. Thus, recorded examinations performed by a radiographer allow the radiologist to devote more of his or her time to more complex tasks. In a situation with limited healthcare resources, this is likely to result in lower costs or more efficient use of available resources, provided that the image quality can be maintained at an adequate level. In most areas of radiology today, standardized methods and predetermined protocols are used. For barium enema, Crawley *et al.* (7) claim that radiographers can be trained to perform standardized examinations from a predetermined protocol without describing pathology. This may result in a development of the radiographers' professional role as well as increased availability of the radiologist for more complex exams. Detailed organ schemes for ultrasound examination with cine documentation, combined with the radiographers' knowledge of abdominal anatomy, seem to make radiographers equally suited to perform abdominal ultrasound investigations.

Documentation of ultrasound examinations is often made only with static images saved from regions of particular interest or where pathology is seen (8). Although the patient may have been examined in a systematic way, only the examiner knows what was seen before and after the static images. The value of re-evaluating static images is thus very limited, and the experience and education of the examiner are crucial with the traditional method (9, 10). With a standardized method, the value of an examination performed by an examiner with limited experience may increase by review by an experienced observer. A great advantage of the standardized method is the possibility for several observers to re-evaluate the dynamic clips at a workstation. This may be an important tool for increasing patient safety.

One obvious limitation of our study is the low number of patients, making the confidence intervals wide and the results less certain. It is possible that a larger material might give more reliable results. It is also important to ascertain whether the results found here with one radiographer and two radiologists who have worked closely together for several years can be generalized to a different setting.

In conclusion, this pilot study showed a satisfactory agreement between the new standardized method and the traditional. If these results are confirmed by larger studies in the future, ultrasound examinations performed by a radiographer and analyzed off-line by a radiologist may be a clinically useful alternative to bedside examinations by a radiologist.

Conflict of interest: None.

REFERENCES

1 Leslie A, Lockyer H, Virjee JP. Who should be performing routine abdominal ultrasound? A prospective double-blind study comparing the accuracy of radiologist and radiographer. *Clin Radiol* 2000;**55**:606–9
2 Lo RHG, Chan PP, Chan LP, *et al.* Routine abdominal and pelvic ultrasound examinations: an audit comparing radiographers and radiologist. *Ann Acad Med Singapore* 2003;**32**:126–8
3 Thorelius L. Can ultrasound benefit from the rest of radiology? 2007. See http://www.auntminnie.com/redirect/redirect.asp?itemid=74965
4 Altman DG. *Practical Statistics for Medical Research*. London: Chapman & Hall, 1991
5 Feinstein AR, Ciccihetti DV. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990;**43**:551–8
6 Viera AJ, Garett JM. Understanding interobserver agreement: The kappa statistic. *Fam Med* 2005;**37**:360–3
7 Crawley M T, Shine B, Booth A. Radiation dose and diagnosticity of barium enema examinations by radiographers and radiologists: a comparative study. *Br J Radiol* 1998;**71**:399–405
8 Tegnander E, Eik-Nes SH. The examiner's ultrasound experience has a significant impact on the detection rate of congenital heart defects at the second-trimester fetal examination. *Ultrasound Obstet Gynecol* 2006;**28**:8–14
9 Pallan M, Linnane J, Ramaia S. Evaluation of an independent, radiographer-led community diagnostic ultrasound service provided to general practitioners. *J Public Health (Oxf)* 2005;**27**:176–81
10 Rosen MP, Levine D, Carpenter JM, *et al.* Diagnostic accuracy with US: remote radiologists versus on-site radiologists' interpretations. *Radiology* 1999;**210**:733–6