

## Inter-rater Agreement and Reliability among General Practitioners and Radiologists on Ultrasound Examinations after Specialized Training Program

This article was published in the following Scient Open Access Journal:

Journal of Primary Health Care and General Practice

Received February 12, 2019; Accepted February 27, 2019; Published March 04, 2019

Mònica Solanes Cabús<sup>\*1</sup>, Jesús Pujol Salud<sup>2</sup>,  
Jesús Alonso Aliaga<sup>3</sup>, Francesc Mayench  
Ariza<sup>4</sup>, Josep Ramon Marsal Mora<sup>5</sup>, Anna  
Sabaté Navarro<sup>6</sup>, Sergi Gòdia López<sup>7</sup>, Carme  
Florensa Piró<sup>8</sup>, Mireia Abellana Sangrà<sup>9</sup>,  
Mireia Marin Muñoz<sup>10</sup>, Laia Llorc Samsó<sup>11</sup>,  
Ramon Mur Garcés<sup>12</sup>, Jose Maria Palacín  
Peruga<sup>13</sup>, Montse Ribes Ribes<sup>14</sup>, Marta Ortega  
Bravo<sup>15</sup>

<sup>\*1</sup>Onze de Setembre Primary Care Center, Institut Català de la Salut (ICS), 25005 Lleida, Spain.

<sup>2</sup>Universitat de Lleida (UdL). Biomedical Research Institute (IRB Lleida), Balaguer Primary Care Center, Institut Català de la Salut (ICS), 25600 Lleida, Spain.

<sup>3</sup>Cervera Primary Care Center, Institut Català de la Salut (ICS), 25200 Lleida, Spain.

<sup>4</sup>Àrea de Gestió Administrativa, Direcció d'Atenció Primària Lleida, Institut Català de la Salut (ICS), 25007 Lleida, Spain.

<sup>5</sup>Josep Ramon Marsal Mora - Institut Universitari d'Investigació en Atenció Primària IDIAP "Jordi Gol i Gurina", 08007 Barcelona, Spain.

<sup>6</sup>Rambla de Ferran Primary Care Center, Institut Català de la Salut (ICS), 25007 Lleida, Spain.

<sup>7</sup>Alcarràs Primary Care Center, Institut Català de la Salut (ICS), Lleida, 25180 Spain.

<sup>8</sup>Borges Blanques Primary Care Center, Institut Català de la Salut (ICS), 25400 Lleida, Spain.

<sup>9</sup>Mollerussa Primary Care Center, Institut Català de la Salut (ICS), 25230 Lleida, Spain.

<sup>10</sup>Alcarràs Primary Care Center, Institut Català de la Salut (ICS), 25180 Lleida, Spain.

<sup>11</sup>Laia Llorc Samsó - Borges Blanques Primary Care Center, Institut Català de la Salut (ICS), 25400 Lleida, Spain.

<sup>12</sup>Eixample Primary Care Center, Institut Català de la Salut (ICS), 25006 Lleida, Spain.

<sup>13</sup>Onze de Setembre Primary Care Center, Institut Català de la Salut (ICS), 25005 Lleida, Spain.

<sup>14</sup>Montse Ribes Ribes - Bordeta Primary Care Center, Institut Català de la Salut (ICS), 25001 Lleida, Spain.

<sup>15</sup>Institut Universitari d'Investigació en Atenció Primària IDIAP "Jordi Gol i Gurina", 08007 Barcelona, Spain. Cappont Primary Care Center, Institut Català de la Salut (ICS), 25001 Lleida, Spain

**\*Corresponding Author:** Mònica Solanes Cabús,  
Onze de Setembre Primary Care Center, Institut  
Català de la Salut (ICS), 25005 Lleida, Spain,  
Email: msolanes.lleida.ics@gencat.cat

### Abstract

**Introduction:** Ultrasound examination is a medical diagnostic test broadly used, due to its properties of being a non-invasive method and its low operational costs. Commonly, Radiologists perform ultrasound examinations, as a part of their medical specialty. However, ultrasound examinations could be applied by other specialties, as a first screening test. For that, a Training program for General Practitioners (GPs) led by Radiologists was designed, and the inter-rate agreement and reliability among GPs and Radiologists was analyzed in order to implement ultrasound examination performed by GPs at Primary Care Centers (PCC). Also, the level of patient' confidence on getting an ultrasound examination done by GPs, instead Radiologist, was measured by a satisfaction survey.

**Methods:** 60-hours course of theoretical and practical training was done, followed by 24-weeks practical sessions at the Radiology Service. Thirteen GPs were enrolled on the Training program. On practical sessions, both GPs and Radiologists assessed each patient and recorded the data on a newly built electronic form. Kappa index was calculated including the overall inter-rater agreement among GPs and Radiologists examinations. After Training program completion, GPs used ultrasound examination in their Primary Care Centers and patient's satisfaction survey was done.

**Results:** The total number of patients assessed by both professionals was 1,928: 1,377 patients had abdominal ultrasound scan, 345 patients had urological ultra Sound scan, and 206 patients had thyroid ultrasound scan. A high degree of agreement among raters was observed ( $\kappa \geq 0.8$ ) in most of the double-checked organs. The patients overall assessment rate was very satisfied: 4.8 out of 5.

**Conclusions:** This study provides insights into ultrasound examination used by GPs in PCC, after an appropriate training. That could be used as fundamental data in improving health care protocols in the near future.

**Keywords:** Ultrasound, General Practitioner, Radiologist, Primary Care, Training, Guidelines, Inter-Rater Agreement, Reliability

### Introduction

Ultrasound examination is a rapid, non-invasive method, and radiation free investigation with a wide variety of clinical applications. It is a medical test used for diagnosis, due to its properties of harmlessness, good tolerance, dynamics and low cost [1]. Traditionally, ultrasound examinations have been a part of the Radiology specialty. However, in recent years, other medical specialties have adapted the use of ultrasound examinations for their patients. This way of conducting ultrasound examinations is known as a point-of-care ultrasound examination [2]. Since long time ago, a few publications looked into the possibilities of utilizing ultrasound examinations in PCC [3]. Interestingly, there are a few studies where General Practitioners (GPs) of Primary Care Centers (PCC) had been trained and they use point-of-care ultrasound examinations daily [4-10].

In Spain, Radiologists have been the medical professionals who used to perform requested ultrasound examinations in Primary Care Centers (PCC). On 2012, the Spanish Ministry of Health recommended to incorporate ultrasound examination in PCC [11]. Gradually, PCC has been equipped with ultrasound machines and training for GPs has been provided. However, a structured education and training on ultrasound

technology is mandatory to keep skills at a sufficient level over time for GPs.

An ultrasound examination-training program for GPs was designed by the Primary Care Management in Lleida (DAP Lleida), and aligned within the current framework of the Catalan government Health plan (2011-2015) [12]. For that, the recommendations of the Royal College of Radiologists in London and the European Federation of Societies for Ultrasound in Medicine and Biology were reviewed and adapted [13,14], and a specific training program for GPs led by Radiologists was organized to educate them on handling of ultrasound machines and processing of the obtained data for clinical diagnosis at PCC. Thirteen GPs were enrolled on the full Training program, where 1,928 patients were assessed. The aim of the current study was to evaluate the quality of the ultrasound examinations performed by GPs after completion of the training program, and analyze the inter-rater agreement and reliability within GPs and Radiologists in a limited range of ultrasound examinations. Also, the patients' satisfaction with the ultrasound examination in a PCC was measured by a survey after the technique implementation. This is the first big study that validates the use of ultrasound examination by GPs. GPs and Radiologist showed a very high degree of agreement and reliability in the performance of the technique as a first level screening.

## Method

The study was conducted in two public Hospitals in Lleida (Spain), 13 GPs and 2 Radiologists performed all the analyzed ultrasound examinations between October 2012 and September 2015. The study was approved by the Institut d'Investigació en Atenció Primària (IDIAP) Jordi Gol i Gurina Ethical Committee (registration number p18/077) and it complied with the Declaration of Helsinki.

## GPs Training Program Description

General Practitioners performed a 60-hours course of theoretical

and practical training; the National Health System Continued Training Committee certified the course. The training included theoretical training on the basic acquisition and interpretation of the images. Later, 13 GPs went for a practical training at the Radiology Service of Hospital de Santa Maria (HSM) and Hospital Universitari Arnau de Vilanova (HUAV) in Lleida (Spain), one day per week through 24-weeks. GPs performed 12 ultrasound examinations per day; results were recorded on an electronic form, and a Radiologist performed the ultrasound examination of the same patients afterwards, he also recorded the results on the same electronic form. At the end of the training program, each GP had carried out an average of 200 ultrasound examinations.

## Data Collection

Ultrasound examinations were performed with an ultrasound device. The ultrasound equipment had a Digital Imaging and Communication in Medicine (DICOM) license that allowed GPs and Radiologists to save all the recorded pictures in the system, allowing them to be double check for training purposes.

Based on the Ultrasound Training Competency Assessment Sheet model [14], an electronic form was generated, where both professionals could record the obtained data in each ultrasound examination. After patient examination, each professional had to classify it as 'normal' (or benign entities) or 'abnormal' (presence of pathological signs). The Radiologist diagnose was taken as a correct option in all the cases, because GPs were still under training. The designed electronic form contributed to the reliability of data collection and to the study's accuracy.

The total number of patients who went under ultrasound examination was 1,928: 1,377 patients had abdominal ultrasound scan (bile duct, gall bladder, liver, pancreas, aorta, kidneys, spleen, bladder), 345 patients had urological ultrasound scan (kidneys, bladder and prostate in men), and 206 patients had thyroid ultrasound scan (Figure 1). Gynecological/obstetric ultrasound examinations, oncology-related patients and advance liver disease patients were excluded from Patients cohort.

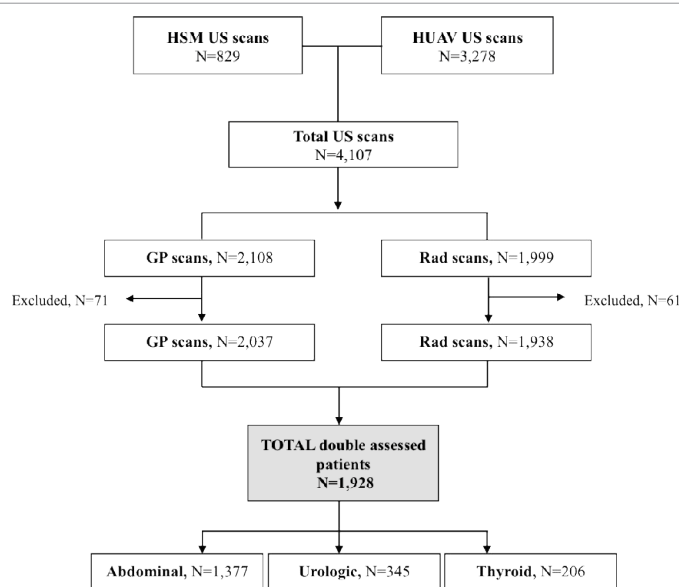


Figure 1: Description of Patient' ultrasound examinations Cohort. US: ultrasound examination. HSM: Hospital de Santa Maria. HUAV: Hospital Universitari Arnau de Vilanova. GP: General Practitioner. Rad: Radiologist

Once the 13 GPs finalized the Training Program, they used ultrasound examination in their medical consultation offices at Primary Care Centers (CAP Balaguer, CAP Cervera, CAP Capped). When an ultrasound examination was necessary, it was performed by a GP and afterwards patients were asked to fill out a short written survey on patient's satisfaction. The survey was anonymous and it was composed by six-questions. The patients could grade the medical service in the consultation scoring from 1 (lowest score, not satisfied) to 5 (highest score, fully satisfied): 1) Was the waiting time for the ultrasound examination adequate? ; 2) Rate the comfort experienced during the ultrasound scan; 3) Were you properly informed about ultrasound examination? ; 4) Rate the level of confidence that an ultrasound is done by GP; 5) Rate the care received from the GP; 6) Rate the overall assessment of the consultation. 170 patients gave feedback on the consultation about the potential inclusion of ultrasound examination performed by GPs at the PCC.

### Statistical analysis

For statistical evaluation, Kappa statistic index values were calculated including the overall inter-rater agreement between the results from the GPs and Radiologists examinations. Measurement of the extent to which data collectors (raters) assign the same score to the same variable is called inter-rater reliability. Kappa value was calculated on each ultrasound examination performed by GPs and Radiologists. Kappa index can be interpreted as follows:  $0.61 \leq \kappa < 0.80$  as substantial/'good' degree of concordance, and  $0.81 \leq \kappa < 1.00$  as almost perfect agreement/'very good' degree of concordance. The risk ratio, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were studied concurrently. The Radiologists assessment was considered the correct one. All patient data was anonymous.

SPSS 15.0 software was used for data management, statistical analysis and plotting of the results. Statistical difference was considered significant at p-values  $\leq 0.05$ .

## Results

### Description of Patients Cohort

After the attendance to the Ultrasound Training program, thirteen GPs had enrolled to a practical 24 weeks program at the Radiology Service of two public Hospitals (Hospital de Santa Maria -HSM; Hospital Universitari Arnau de Vilanova -HUAV). The total number of ultrasound examinations was 4,107: 3,289 examinations were performed at the HUAV and 829 examinations at the HSM. More precisely, GPs performed a total number of 2,037 examinations (71 were excluded) and Radiologist performed 1,938 examinations (61 were excluded). The double assessed patient cohort was established with a total of 1,928 ultrasound examinations. Ultrasound examinations were classified on 1,377 abdominal, 345 urological and 206 thyroid ultrasound examinations (Figure 1).

### Concordance on Ultrasound examinations among GPs and Radiologists

The degree of concordance among GPs and Radiologists ultrasound examinations was *good* in pancreas ( $\kappa=0.629$ ), thyroids ( $\kappa=0.749$ ), left kidney ( $\kappa=0.792$ ), right kidney ( $\kappa=0.780$ ), and bladder ( $\kappa=0.789$ ). The concordance was *very good* in bile

duct ( $\kappa=0.821$ ), gall bladder ( $\kappa=0.824$ ), liver ( $\kappa=0.825$ ), and spleen ( $\kappa=0.876$ ). And it reached an almost perfect agreement in the aorta examination ( $\kappa=0.949$ ) (Table 1).

The sensitivity of GPs examinations (defined as the percentage of positive correct identified) was *very high* in bile duct (97.7%), spleen (98.8%) and aorta (99.5%); *high* in liver (93.2%), gall bladder (92.5%), pancreas (93.7%), right kidney (89.7%), left kidney (89.3%) and bladder (98.1%). However, the lowest sensitivity was at thyroids (70.2%).

Precision and accuracy of the Training program in Ultrasound examinations was measured by the Positive/Negative predictive values (PPV/NPV). PPV was *very high* in pancreas (99.8%), aorta (99.8%) and bile duct (99.5%). A bit lower percentage was showed by left kidney (95.5%) and thyroid (95.2%). So, the Training program was adequate and validated.

However, some divergence indicators were also analyzed to know the extent of agreement among GPs and Radiologist reliability. For that, the percentage of patients scored as 'abnormal' by the Radiologist and 'normal' by the GP (RADA) was calculated and it was very low (range 0.1% - 2.8%), and the risk ratio was also very low in all the examined organs (Table 2). Moreover, pancreas showed the lowest concordance index and 6% of GPs examinations were not confirmed by the Radiologist. Supplementary data is provided on divergences among GPs and Radiologist for each organ (see Supplementary data) (Tables 3-11).

### Patient satisfaction survey

170 patients filled out a survey where they answered anonymously six questions. The overall assessment rating was very high ( $4.8 \pm 0.5$ ), and the rate of confidence for an ultrasound performed by GPs was also really high ( $4.8 \pm 0.5$ ). The survey showed direct feedback on the Training program and it was a legitimate indicator for improving the services and strategic goals in healthcare organizations (Figure 2).

## Discussion

This study analyzed the inter-rater agreement and reliability among General Practitioners (GPs) and Radiologist ultrasound examinations after specialized training program for GPs to potentially establish ultrasound examinations performed by GPs at Primary Care Centers (PCC), as a screening technique in medical consultations. Both GPs and Radiologist assessed a total of 1,928 examinations, during the practical phase of the training program. Ultrasound examinations were classified on 1,377 abdominal, 345 urological and 206 thyroid ultrasound examinations. A high

Table 1: Reasons for requesting ultrasound examinations

	N	%	95%CI
Other diagnoses	1,801	27.8%	26.7% - 28.9%
Processes of urological origin	1,514	23.3%	22.3% - 24.4%
Processes with abdominal pain	1,044	16.1%	15.2% - 17%
Processes of hepatic origin	607	9.4%	8.7% - 10.1%
Analytical alteration	567	8.7%	8.1% - 9.4%
Thyroid gland processes	488	7.5%	6.9% - 8.2%
General medical check-up without symptoms	463	7.1%	6.5% - 7.8%
Total*	6,484		

(\*): 296 other diagnoses, none of them with a prevalence greater than 1%.

**Table 2:** Concordance on ultrasound examinations among General Practitioners and Radiologists.

Ultrasound assessment	N	RADn <sup>a</sup> n (%)	GPn <sup>b</sup> n (%)	κ 95%CI	Not match n (%)	GPa <sup>c</sup> n (%)	RADa <sup>d</sup> n (%)	RR <sup>e</sup>	Sens <sup>f</sup> (%)	Spec <sup>g</sup> (%)	PPV <sup>h</sup> (%)	NPV <sup>i</sup> (%)
Abdominal												
Liver	1,377	965 (70.1)	936 (68.0)	0.825 (0.77-0.88)	103 (7.5)	66 (4.8)	37 (2.7)	1.8	93.2	91.0	96.0	85.0
Bile duct	1,377	1,280 (93.0)	1,256 (91.2)	0.821 (0.77-0.87)	36 (2.6)	30 (2.2)	6 (0.4)	5.0	97.7	93.8	99.5	75.2
Gall bladder	1,377	979 (71.1)	936 (68.0)	0.824 (0.77-0.88)	103 (7.5)	73 (5.3)	30 (2.2)	2.4	92.5	92.5	96.8	83.4
Pancreas	1,377	1,294 (94.0)	1,214 (88.2)	0.629 (0.58-0.68)	84 (6.1)	82 (6.0)	2 (0.1)	41.0	93.7	97.6	99.8	49.7
Spleen	1,377	1,249 (90.7)	1,248 (90.6)	0.876 (0.82-0.93)	29 (2.1)	15 (1.1)	14 (1.0)	1.1	98.8	89.1	98.9	88.4
Aorta	1,377	1,296 (94.1)	1,292 (93.8)	0.949 (0.90-0.99)	8 (0.6)	6 (0.4)	2 (0.1)	3.0	99.5	97.5	99.8	92.9
Urological												
Bladder	1,722	1,589 (92.3)	1,582 (91.9)	0.789 (0.74-0.84)	53 (3.1)	30 (1.7)	23 (1.3)	1.3	98.1	82.7	98.5	78.6
Right kidney	1,722	1,157 (67.2)	1,082 (62.8)	0.792 (0.75-0.84)	163 (9.5)	119 (6.9)	44 (2.6)	2.7	89.7	92.2	95.9	81.4
Left kidney	1,722	1,157 (67.2)	1,082 (62.8)	0.780 (0.73-0.83)	173 (10.0)	124 (7.2)	49 (2.8)	2.5	89.3	91.3	95.5	80.6
Thyroids	206	57 (27.7)	42 (20.4)	0.749 (0.62-0.88)	19 (9.2)	17 (8.3)	2 (1.0)	8.5	70.2	98.7	95.2	89.6

**N:** number of examinations. **κ:** kappa index. **a) RADn:** According to the Radiologist, percentage of patients scored as 'normal'. **b) GPn:** According to the General Practitioner, percentage of patients scored as 'normal'. **Not match:** Percentage of patients of without concordance in the score given by the Radiologist and the GP. **c) GPa:** Percentage of patients scored as 'abnormal' by the GP and 'normal' by the Radiologist. **d) RADa:** Percentage of patients scored as 'abnormal' by the Radiologist and 'normal' by the GP. **e) RR:** Risk Ratio of GPa and RADa; RR>1 indicates a higher probability of GP scores 'abnormal'. **f) Sens:** Sensitivity. **g) Spec:** Specificity. **h) PPV:** Positive Predictive Value. **i) NPV:** Negative Predictive Value.

**Table 3:**

LIVER	GP		RAD		Non-concordance	
	N	n (%)	N	n (%)	N	n (%)
Liver LFU injury - Solid hyperechoic nodule	103	1 (1%)	103	1 (1%)	103	2 (1.9%)
Liver Size - Normal	103	99 (96.1%)	103	101 (98.1%)	103	6 (5.8%)
Liver Size - Hepatomegaly	103	4 (3.9%)	103	1 (1%)	103	5 (4.9%)
Liver Injury - Without injury	103	84 (81.6%)	103	80 (77.7%)	103	42 (40.8%)
Liver LFU injury - Probable Haemangioma	103	1 (1%)	103	103 (0%)	103	1 (1%)
Liver LM Injury - Probable haemangioma	103	3 (2.9%)	103	103 (0%)	103	3 (2.9%)
Liver Vascular - Portal vein	103	30 (29.1%)	103	103 (0%)	103	30 (29.1%)
Liver Eco-structure- Homogeneous	103	92 (89.3%)	103	96 (93.2%)	103	18 (17.5%)
Liver Eco-structure- Heterogeneous	103	11 (10.7%)	103	6 (5.8%)	103	17 (16.5%)
Liver Stenosis - No	103	80 (77.7%)	103	94 (91.3%)	103	32 (31.1%)
Liver Stenosis - Yes	103	23 (22.3%)	103	8 (7.8%)	103	31 (30.1%)
Liver LFU injury - Significant calcification	103	4 (3.9%)	103	6 (5.8%)	103	10 (9.7%)
Liver LM Injury - Significant calcifications	103	1 (1%)	103	1 (1%)	103	2 (1.9%)
Liver LFU injury - Simple cystic	103	6 (5.8%)	103	6 (5.8%)	103	12 (11.7%)
Liver LM Injury - Simple cystic	103	3 (2.9%)	103	9 (8.7%)	103	12 (11.7%)
Liver LM Injury - Probable focal steatosis	103	1 (1%)	103	0 (0%)	103	1 (1%)
Liver Location - Right hepatic lobe	103	18 (17.5%)	103	16 (15.5%)	103	34 (33%)
Liver Location - Left hepatic lobe	103	4 (3.9%)	103	11 (10.7%)	103	15 (14.6%)
Liver Vascular - Suprahepatic	103	25 (24.3%)	103	103 (0%)	103	25 (24.3%)

**Table 4:**

BILE DUCT	MF		RAD		Non-concordance	
	N	n (%)	N	n (%)	N	n (%)
Bile duct Art. Hepatic - Dilatation	36	7 (19.4%)	36	1 (2.8%)	36	8 (22.2%)
Bile duct Extra-hepatic - Cholelith normal	36	31 (86.1%)	36	34 (94.4%)	36	7 (19.4%)
Bile duct Extra-hepatic - Cholelith dilated	36	4 (11.1%)	36	0 (0%)	36	4 (11.1%)
Bile duct Intra-hepatic - No	36	31 (86.1%)	36	32 (88.9%)	36	9 (25%)
Bile duct Extra-hepatic - Post cholecystectomy normal	36	1 (2.8%)	36	0 (0%)	36	1 (2.8%)
Bile duct Aerobilia - No	36	31 (86.1%)	36	33 (91.7%)	36	8 (22.2%)
Bile duct Aerobilia - Yes	36	2 (5.6%)	36	1 (2.8%)	36	3 (8.3%)
Bile duct Art. Hepatic - Dilatation	36	7 (19.4%)	36	1 (2.8%)	36	8 (22.2%)

Table 5:

GALL BLADDER	MF		RAD		Non-concordance	
	N	n (%)	N	n (%)	N	n (%)
Gall bladder - Contents - Neoformation	103	3 (2.9%)	103	3 (2.9%)	103	6 (5.8%)
Gall bladder - Walls - Normal	103	64 (62.1%)	103	84 (81.6%)	103	58 (56.3%)
Gall bladder - Walls - Thick	103	10 (9.7%)	103	7 (6.8%)	103	17 (16.5%)
Gall bladder - Walls - Focal wall injury	103	4 (3.9%)	103	1 (1%)	103	5 (4.9%)
Gall bladder - Walls - Multiple focal injury	103	1 (1%)	103	1 (1%)	103	2 (1.9%)
Gall bladder - Contents - Lithiasis	103	27 (26.2%)	103	10 (9.7%)	103	37 (35.9%)
Gall bladder - Contents - Without	103	47 (45.6%)	103	80 (77.7%)	103	79 (76.7%)
Gall bladder - Contents - Lot	103	8 (7.8%)	103	1 (1%)	103	9 (8.7%)

Table 6:

RIGHT KIDNEY	MF		RAD		Non-concordance	
	N	n (%)	N	n (%)	N	n (%)
Right kidney - Dilatation - Yes	163	18 (11%)	163	8 (4.9%)	163	26 (16%)
Right kidney - LU Injury - Solid injury	163	1 (0.6%)	163	2 (1.2%)	163	3 (1.8%)
Right kidney - Size - Normal	163	143 (87.7%)	163	158 (96.9%)	163	25 (15.3%)
Right kidney - Size - Large	163	4 (2.5%)	163	2 (1.2%)	163	6 (3.7%)
Right kidney - Size - Small	163	13 (8%)	163	2 (1.2%)	163	15 (9.2%)
Right kidney - Morphological alt - Without	163	158 (96.9%)	163	162 (99.4%)	163	6 (3.7%)
Right kidney - Morphological alt - Agenesis	163	1 (0.6%)	0	163 (0%)	163	1 (0.6%)
Right kidney - Lithiasis - No	163	148 (90.8%)	163	157 (96.3%)	163	21 (12.9%)
Right kidney - Lithiasis - Yes	163	12 (7.4%)	163	5 (3.1%)	163	17 (10.4%)
Right kidney - Dilatation - No	163	142 (87.1%)	163	154 (94.5%)	163	30 (18.4%)
Right kidney - LU Injury - Atypical cyst	163	1 (0.6%)	163	1 (0.6%)	163	2 (1.2%)
Right kidney Normal	163	44 (27%)	0	163 (0%)	163	44 (27%)
Right kidney - Eco-structure- Normal	163	158 (96.9%)	163	160 (98.2%)	163	8 (4.9%)
Right kidney - Eco-structure- Echogenic	163	2 (1.2%)	163	2 (1.2%)	163	4 (2.5%)
Right kidney - Parenchyma thickness - Normal	163	154 (94.5%)	163	159 (97.5%)	163	13 (8%)
Right kidney - Parenchyma thickness - Reduced	163	6 (3.7%)	163	3 (1.8%)	163	9 (5.5%)
Right kidney - Parenchyma scar - Without	163	158 (96.9%)	163	161 (98.8%)	163	7 (4.3%)
Right kidney - Parenchyma scar - Unique	163	2 (1.2%)	163	1 (0.6%)	163	3 (1.8%)
Right kidney - Morphological alt - Renal Ectopic	163	1 (0.6%)	0	163 (0%)	163	1 (0.6%)
Right kidney - Injury - Without injury	163	95 (58.3%)	163	131 (80.4%)	163	100 (61.3%)
Right kidney - LU Injury - Simple cortical cyst	163	32 (19.6%)	163	11 (6.7%)	163	43 (26.4%)
Right kidney - LU Injury - Simple sinusal cyst	163	9 (5.5%)	163	2 (1.2%)	163	11 (6.7%)
Right kidney - LM Injury - Simple cortical cyst	163	18 (11%)	163	10 (6.1%)	163	28 (17.2%)
Right kidney - LM Injury - Simple sinusal cyst	163	14 (8.6%)	163	10 (6.1%)	163	24 (14.7%)

degree of agreement among raters was shown ( $\kappa \geq 0.8$ ) in most of the double-checked organs. The patients overall assessment rate was very satisfied ( $4.8 \pm 0.5$ ) after the ultrasound examination was done by a trained GP in a PCC. The present research results can be utilized as reliable data to define effective and new protocols in Primary Care in the near future.

In the current study, the total number of double assessed patients (N= 1,928) was higher than in previous studies, where a much lower number of patients were double-checked by a GP and a Radiologist (N=120, [15]; N=114, [2]). This characteristic gave stronger statistical power and relevance to our current data (Figure 1).

Interestingly, a high sensitivity was observed (89%) in most of the double-checked organs (Table 1). The lowest concordance was found in pancreas ( $\kappa=0.629$ ; NPV=50%) and thyroid gland ( $\kappa=0.749$ ; Sen=70.2%). This could be due to a lack of specific training in those organs, where the physiological complexity is

intrinsically high. However, pancreas inter-rater agreement in the current study was higher than in previous ones ( $\kappa=0.38$ ; [15]).

On the other hand, on abdominal/urological areas a similar concordance levels were observed when the ultrasound examination was done by GPs ( $\kappa=0.93$ , [16];  $\kappa=0.73$ , [17]). Regarding ascites, abdominal aortic aneurism and gallstones, the consistency was almost 100% [6]. However it has shown that the ultrasound examination could modify the clinical treatment of abdominal pain in up to 64% of cases, reducing referrals to other specialists up to 40% [18]. In addition, when a Training Program was focused in one single organ/disease (Abdominal Aortic Aneurysm -AAA), the degree of concordance was perfect ( $\kappa=1$ , [19]) and it had high diagnostic validity [20]. The concordance in the diagnosis of profound venous thrombosis between GPs and vascular experts was very good ( $\kappa=0.86$ , [21]); and on hypertensive patients, the concordance degree was 'very high' ( $\kappa > 0.83$ , [22]).

Table 7:

LEFT KIDNEY	MF		RAD		Non-concordance	
	N	n (%)	N	n (%)	N	n (%)
Left kidney - Dilatation of duct - Yes	173	14 (8.1%)	173	4 (2.3%)	173	18 (10.4%)
Left kidney - LU - Solid injury	173	2 (1.2%)	173	1 (0.6%)	173	3 (1.7%)
Left kidney - Size - Normal	173	154 (89%)	173	170 (98.3%)	173	22 (12.7%)
Left kidney - Size - Large	173	4 (2.3%)	173	3 (1.7%)	173	7 (4%)
Left kidney - Size - Small	173	13 (7.5%)	173	0 (0%)	173	13 (7.5%)
Left kidney - Morphological alt - Without	173	166 (96%)	173	172 (99.4%)	173	8 (4.6%)
Left kidney - Morphological alt - Agenesis	173	1 (0.6%)	173	0 (0%)	173	1 (0.6%)
Left kidney - Lithiasis - No	173	163 (94.2%)	173	167 (96.5%)	173	16 (9.2%)
Left kidney - Lithiasis - Yes	173	8 (4.6%)	173	6 (3.5%)	173	14 (8.1%)
Left kidney - Dilatation of duct - No	173	157 (90.8%)	173	169 (97.7%)	173	20 (11.6%)
Left kidney - LU - Atypical cyst	173	1 (0.6%)	173	0 (0%)	173	1 (0.6%)
Left kidney - Normal	173	49 (28.3%)	173	0 (0%)	173	49 (28.3%)
Left kidney - Eco-structure- Normal	173	170 (98.3%)	173	171 (98.8%)	173	5 (2.9%)
Left kidney - Eco-structure- Echogenic	173	1 (0.6%)	173	2 (1.2%)	173	3 (1.7%)
Left kidney - Parenchyma thickness - Normal	173	161 (93.1%)	173	170 (98.3%)	173	15 (8.7%)
Left kidney - Parenchyma thickness - Reduced	173	10 (5.8%)	173	3 (1.7%)	173	13 (7.5%)
Left kidney - Parenchyma scar - Without	173	169 (97.7%)	173	173 (100%)	173	4 (2.3%)
Left kidney - Parenchyma scar - Unique	173	2 (1.2%)	173	0 (0%)	173	2 (1.2%)
Left kidney - Morphological alt - Double system	173	4 (2.3%)	173	1 (0.6%)	173	5 (2.9%)
Left kidney - Injury - Without injury	173	97 (56.1%)	173	135 (78%)	173	114 (65.9%)
Left kidney - LU Injury - Simple cortical cyst	173	27 (15.6%)	173	7 (4%)	173	34 (19.7%)
Left kidney - LU - Simple sinusal cyst	173	10 (5.8%)	173	2 (1.2%)	173	12 (6.9%)
Left kidney - LU - Pyramidal hypertrophy	173	0 (0%)	173	1 (0.6%)	173	1 (0.6%)
Left kidney - LM - Simple cortical cyst	173	26 (15%)	173	15 (8.7%)	173	41 (23.7%)
Left kidney - LM - Simple sinusal cyst	173	18 (10.4%)	173	23 (13.3%)	173	41 (23.7%)

Table 8

SPLEEN	MF		RAD		Non-concordance	
	N	n (%)	N	n (%)	N	n (%)
Spleen - LU - Solid injury	29	1 (3.4%)	0	0 (0%)	29	1 (3.4%)
Spleen - Size - Normal	29	23 (79.3%)	29	16 (55.2%)	29	19 (65.5%)
Spleen - Size - Megalia	29	5 (17.2%)	29	11 (37.9%)	29	16 (55.2%)
Spleen - Eco-structure- Homogeneous	29	28 (96.6%)	29	26 (89.7%)	29	4 (13.8%)
Spleen - Injury - Without injury	29	24 (82.8%)	29	26 (89.7%)	29	8 (27.6%)
Spleen - LM - Granuloma	29	1 (3.4%)	29	0 (0%)	29	1 (3.4%)
Spleen - LU - Cyst	29	2 (6.9%)	29	0 (0%)	29	2 (6.9%)
Spleen - Splenectomy	29	1 (3.4%)	29	1 (3.4%)	29	2 (6.9%)

Table 9

AORTA	MF		RAD		Non-concordance	
	N	n (%)	N	n (%)	N	n (%)
Aorta - Size - Dilatation	8	4 (50%)	0	0 (0%)	8	4 (50%)
Aorta - Size - Normal	8	4 (50%)	8	6 (75%)	8	6 (75%)

Table 10

BLADDER	MF		RAD		Non-concordance	
	N	n (%)	N	n (%)	N	n (%)
Bladder - Walls - Focal wall injury (polyp)	53	3 (5.7%)	53	3 (5.7%)	53	6 (11.3%)
Bladder - Contents - Neoformation	53	3 (5.7%)	53	3 (5.7%)	0	0 (0%)
Bladder - Walls - Normal	53	44 (83%)	53	45 (84.9%)	53	17 (32.1%)
Bladder - Walls - Thick	53	7 (13.2%)	53	4 (7.5%)	53	11 (20.8%)
Bladder - Contents - Lithiasis	53	3 (5.7%)	53	1 (1.9%)	53	4 (7.5%)

Table 11:

PANCREAS	MF		RAD		Non-concordance	
	N	n (%)	N	n (%)	N	n (%)
Pancreas Partially Pathological	84	1 (1.2%)	0	84 (0%)	84	1 (1.2%)
Pancreas - Not assessable	84	55 (65.5%)	0	84 (0%)	84	55 (65.5%)
Pancreas - Size - Normal	84	28 (33.3%)	84	82 (97.6%)	84	58 (69%)
Pancreas - Size - Large	84	1 (1.2%)	0	84 (0%)	84	1 (1.2%)
Pancreas - Size - Atrophic	84	0 (0%)	84	0 (0%)	84	0 (0%)
Pancreas - Eco-structure- Homogeneous	84	28 (33.3%)	84	82 (97.6%)	84	58 (69%)
Pancreas - Eco-structure- Heterogeneous	84	1 (1.2%)	0	84 (0%)	84	1 (1.2%)
Pancreas - Eco-structure- Calcifications	84	0 (0%)	84	0 (0%)	84	0 (0%)
Pancreas - Injury - Without injury	84	25 (29.8%)	84	82 (97.6%)	84	61 (72.6%)
Pancreas - LU – Solid	84	1 (1.2%)	0	84 (0%)	84	1 (1.2%)
Pancreas - LU – Cystic	84	2 (2.4%)	0	84 (0%)	84	2 (2.4%)
Pancreas - LU - Indeterminate	84	1 (1.2%)	0	84 (0%)	84	1 (1.2%)
Pancreas - LM - Solid	84	0 (0%)	84	0 (0%)	84	0 (0%)
Pancreas - LM - Cystic	84	0 (0%)	84	0 (0%)	84	0 (0%)
Pancreas - LM - Indeterminate	84	0 (0%)	84	0 (0%)	84	0 (0%)
Pancreas - Wirsung - No	84	29 (34.5%)	84	82 (97.6%)	84	57 (67.9%)
Pancreas - Wirsung - Yes	84	0 (0%)	84	0 (0%)	84	0 (0%)

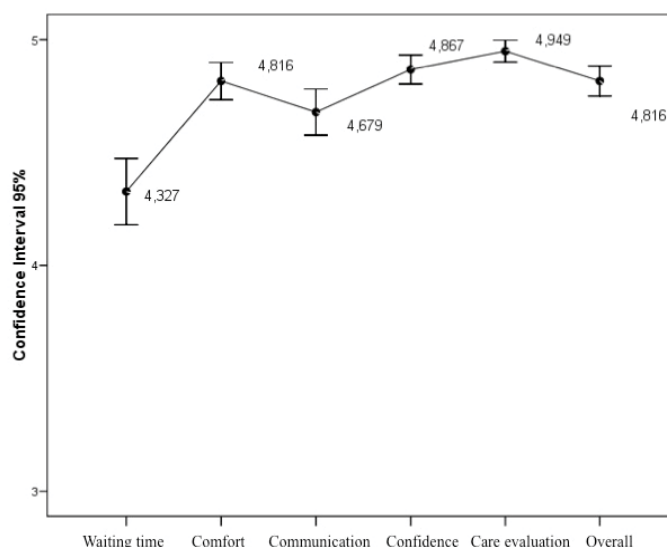


Figure 2: Patient satisfaction survey:

Waiting time: Was the waiting time for the ultrasound examination adequate?

Comfort: Rate the experienced comfort during the ultrasound scan.

Communication: Were you properly informed about ultrasound examination? Confidence: Rate the level of confidence that an ultrasound is done by GP.

Care evaluation: Rate the care received from the GP.

Overall: Rate the overall assessment of the consultation.

The highest discrepancies (GPa, Table 1) were due to the more frequent GPs judgment as 'abnormal', and Radiologist scored as 'normal'. This could be related with the progress of GPs on ultrasound examinations, which generated self-distrust in the initial part and they might mislead some examinations. The difference on academic background and experience was also a variable to take in account, because Radiologists had longer experience and specific training for longer time. So, GPs could not reach the same professional standards only with the practical sessions. However, this would be easy to overcome with time and practical training.

To know the patients perception of ultrasound examination performed by GPs, a survey was designed and anonymously filled out. The new establish approach was successfully accepted in PCC by general public and confirmed by patients' survey that showed high rates of comfort, confidence and care evaluation (Figure 2). In fact, a brief analysis in clinical practice showed that the most prevalent reasons to request an ultrasound examination at the PCC are urological issues (23%), abdominal pain (16%) and hepatic processes (9.4%) (data not shown). It would be very beneficial if GPs could perform, as first screening, ultrasound examinations instead refereeing the patient to the Radiologist in another Hospital.

The current study represents a step forward in the quality of ultrasound examinations performed by GPs. Some recommendations have been done by academies and professionals: structured education with certification, continuous training to keep skills at a sufficient level over time, and choose specific examinations that are more accessible to ensure quality [4, 5, 7, 10, 15]. The sensitivity achieved by GPs appeared suboptimal and ultrasound scan is a reliable tool. Indeed, this would transform GPs ability to achieve a rapid diagnosis for their patients.

GPs implemented ultrasound examination after the training program completion; they have been applying this approach on 70% of the general population attended in a PCC. From 2013 to 2016, they performed a total of 10,946 ultrasound examinations. Currently, the new approach has been established in 12 out of 23 PCC in Lleida state.

## Conclusions

To our knowledge, this is the first study conducted in Spain that investigated inter-rate agreement and reliability among GPs and Radiologists, after a specified Training Program that covers 1928 patients. Particularly, it developed an ultrasound Training program for GPs, and then analyzed the Program rate of success. In most of the double-checked organs, a high degree of agreement among raters and very high sensitivity of GPs examinations were observed. The research results can be utilized as reliable data to consolidate ultrasound scans in PCC performed by experienced and trained GPs, and to implement actual clinical practices.

## Declarations

### Ethics approval and consent to participate

Oral consent was obtained from participants throughout a standard medical consultation. Permission was obtained from the two hospitals involved in the study (Hospital Universitari Arnau de Vilanova and Hospital Universitari de Santa Maria).

The study was approved by the Institut d'Investigació en Atenció Primària (IDIAP) Jordi Gol i Gurina Ethical Committee (registration number p18/077).

### Consent for publication

"Not applicable", as no personal information is provided.

### Availability of data and material

The datasets used are available from the corresponding author on reasonable request.

### Competing interests

The authors declare that they have no competing interests.

### Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

### Acknowledgements

The authors would like to thank all patients who participated in the Study.

This project was supported by the Territorial Health Management

Department of the Institut Català de la Salut (ICS) in Lleida; the biggest two public Hospitals in Lleida state (Hospital Universitari Arnau de Vilanova -HUAV/Hospital de Santa Maria -HSM); the Departament de Salut de Catalunya (CatSalut); and, the Diputació de Lleida. The Primary Care Management (DAP Lleida) has been decisive in the development of this program, due to its complexity.

Special thanks to Dr. Quintí Foguet, for his useful help and advice with the paper writing and proof reading.

We would like to thank the 'Lleida Institut Català de la Salut Computing Systems Unit' to develop and set up the computer interphase where data was recorded for this study.

Special thanks for two Radiologists, who trained the GPs in the practical part of the Training Program, and to Albert Seoane for proof reading the manuscript.

Finally, but not last, we would like to thank the GPs who volunteer and got trained to use the new acquired skills in their jobs to achieve a better clinical practice.

"Lleida ultrasound examination group" is composed by Mònica Solanes Cabús, Jesús Pujol Salud, Jesús Alonso Aliaga, Anna Sabaté Navarro, Sergi Gòdia López, Carme Florensa Piró, Mireia Abellana Sangrà, Mireia Marín Muñoz, Laia Llord Samsó, Ramon Mur Garcés, Jose Maria Palacín Peruga, Montse Ribes Ribes and Marta Ortega Bravo.

## References

1. Valentín-López B, Blasco-Amaro JA. Plan de uso adecuado de tecnologías de diagnóstico por imagen en patología abdominal en atención primaria y especializada.: *Agencia Lái n Entralgo*. 2010.
2. Lindgaard K, Riisgaard L. Validation of ultrasound examinations performed by general practitioners'. *Scandinavian journal of primary health care*. 2017;35(3):256-261.
3. Bratland SZ. Ultrasonic diagnosis in general practice. An evaluation study. *Tidsskr Nor Laegeforen*. 1985 ;105(28):1939-1940.
4. Wordsworth S, Scott A. Ultrasound scanning by general practitioners: is it worthwhile? *Journal of public health medicine*. 2002; 24(2):88-94.
5. Glaso M, Medias IB, Straand J. Diagnostic ultrasound in general practice. *Tidsskrift for den Norske laegeforening*. 2007;127(15):1924-1927.
6. Suramo I, Merikanto J, Paivansalo M, Reinikainen H, Rissanen T, Takalo R. General practitioner's skills to perform limited goal-oriented abdominal US examinations after one month of intensive training. *European journal of ultrasound*. 2002 ;15(3):133-138.
7. Mengel-Jorgensen T, Jensen MB. Variation in the use of point-of-care ultrasound in general practice in various European countries. Results of a survey among experts. *The European journal of general practice*. 2016 ;22(4):274-277.
8. Fakhri AA. Nuclear medicine consultation: A useful tool in primary care to enable more accurate diagnosis. *Journal of family medicine and primary care*. 2017;6(2):201-203.
9. Suman A, Schaafsma FG, van de Ven PM, et al: Effectiveness of a multifaceted implementation strategy compared to usual care on low back pain guideline adherence among general practitioners. *BMC health services research*. 2018 ; 18(1):358.
10. Rikley E, Boillat-Blanco N, Meuwly JY, Breuss E, Senn N: Ultrasonography : an useful diagnostic tool for the general practitioner. *Revue medicale suisse*. 2017;13(562):990-994.
11. Ministerio de Sanidad SSeI. Cartera de servicios comunes del Sistema Nacional de Salud. In.: Gobierno de España. 2012.
12. Catalunya Gd: Pla de salut de Catalunya 2011-2015. In. Edited by Salut Dd. 2012.



13. Ultrasound training recommendations for medical and surgical specialties, Second edition. London: The Royal College of Radiologists. 2012.
14. Ultrasound training recommendations for medical and surgical specialties, Third edition. London The Royal College of Radiologists. 2017.
15. Esquerria M, Roura Poch P, Masat Tico T, Canal V, Maideu Mir J, Cruxent R. Abdominal ultrasound: a diagnostic tool within the reach of general practitioners. *Atencion primaria*. 2012;44(10):576-583.
16. Garcia Burriel L, Berjon Chamorro A, Buron Llamazares JL, Onrubia Baticon C, Nunez A: Can ultrasound become a family doctor tool? Experience in a rural health centre. *Atencion primaria*. 2009;41(11):651-652.
17. Pertierra-Galindo N, Vaquero-Lucas E, Muñoz-Quirós-Aliaga S, Jaraba-Mezquida M, Gamarra-Mendoza N, et al: Grado de acuerdo diagnóstico entre ecografías abdominales realizadas por médicos de familia de un centro de salud madrileño y las derivadas al segundo nivel asistencial. *EuroEco*. 2012 ;3(3):67-73.
18. Speets AM, Hoes AW, van der Graaf Y, et al. Upper abdominal ultrasound in general practice: indications, diagnostic yield and consequences for patient management. *Family practice*. 2006;23(5):507-511.
19. Siso-Almirall A, Gilabert Sole R, Bru Saumell C, et al: Feasibility of hand-held-ultrasonography in the screening of abdominal aortic aneurysms and abdominal aortic atherosclerosis. *Medicina clinica*. 2013;141(10):417-422.
20. Bravo-Merino L, Gonzalez-Lozano N, Maroto-Salmon R, Meijide-Santos G, Suarez-Gil P, Fañanás-Mastral A . Validity of the abdominal ecography in primary care for detection of aorta abdominal aneurism in male between 65 and 75 years. *Atencion primaria* 2017.
21. Mumoli N, Vitale J, Giorgi-Pierfranceschi M, et al. General Practitioner-Performed Compression Ultrasonography for Diagnosis of Deep Vein Thrombosis of the Leg: A Multicenter, Prospective Cohort Study. *Annals of family medicine*. 2017;15(6):535-539.
22. Evangelista L, Juncadella E, Copetti S, Pareja A, Torradabella J, Evangelista A. Diagnostic usefulness of pocket echography performed in hypertensive patients by a general practitioner. *Medicina clinica*. 2013;141(1):1-7.