

Power and Color Doppler Ultrasound Settings for Inflammatory Flow

Impact on Scoring of Disease Activity in Patients With Rheumatoid Arthritis

Søren Torp-Pedersen,¹ Robin Christensen,² Marcin Szkudlarek,³ Karen Ellegaard,² Maria Antonietta D'Agostino,⁴ Annamaria Iagnocco,⁵ Esperanza Naredo,⁶ Peter Balint,⁷ Richard J. Wakefield,⁸ Arendse Torp-Pedersen,² and Lene Terslev⁹

Objective. To determine how settings for power and color Doppler ultrasound sensitivity vary on different high- and intermediate-range ultrasound machines

Supported by the Oak Foundation. Dr. Terslev's work was supported by the Danish Rheumatism Association. Study-related travel to Denmark for some of the authors was funded by General Electric.

¹Søren Torp-Pedersen, MD: Copenhagen University Hospital, Frederiksberg, Denmark, and Copenhagen University Hospital at Glostrup, Glostrup, Denmark; ²Robin Christensen, MSc, PhD (Biostat), Karen Ellegaard, PT, PhD, Arendse Torp-Pedersen, BM: Copenhagen University Hospital, Frederiksberg, Denmark; ³Marcin Szkudlarek, MD, PhD: University of Copenhagen Hospital at Køge, Køge, Denmark; ⁴Maria Antonietta D'Agostino, MD, PhD: Ambroise-Paré Hospital, AP-HP, Boulogne-Billancourt, France; ⁵Annamaria Iagnocco, MD: Sapienza Università di Roma, Rome, Italy; ⁶Esperanza Naredo, PhD: Hospital General Universitario Gregorio Marañón, Madrid, Spain; ⁷Peter Balint, MD: National Institute of Rheumatology and Physiotherapy, Budapest, Hungary; ⁸Richard J. Wakefield, BM, MD, FRCP: University of Leeds and Chapel Allerton Hospital, Leeds, UK; ⁹Lene Terslev, MD, PhD: Copenhagen University Hospital at Glostrup, Glostrup, Denmark.

Dr. S. Torp-Pedersen has received consulting fees, speaking fees, and/or honoraria from General Electric, Esaote, AbbVie, and MSD (less than \$10,000 each). Dr. Szkudlarek has received speaking fees and/or honoraria (less than \$10,000) for lectures at courses sponsored by General Electric. Dr. Ellegaard has received speaking fees and/or honoraria (less than \$10,000) for lectures at courses sponsored by General Electric. Dr. D'Agostino has received consulting fees, speaking fees, and/or honoraria from Bristol-Myers Squibb, AbbVie, and General Electric (less than \$10,000 each). Dr. Iagnocco has received consulting fees from AbbVie and General Electric (less than \$10,000 each). Dr. Naredo has received consulting fees, speaking fees, and/or honoraria from AbbVie, Roche, Bristol-Myers Squibb, Pfizer, UCB, MSD, General Electric, and Esaote (less than \$10,000 each). Dr. Balint has received consulting fees, speaking fees, and/or honoraria from General Electric, Esaote, and SonoSite (less than \$10,000 each) and from Philips (more than \$10,000). Dr. Wakefield has received speaking fees from AbbVie and consulting fees from General Electric (less than \$10,000 each). Dr. Terslev has received speaking fees from General Electric (less than \$10,000).

and to evaluate the impact of these changes on Doppler scoring of inflamed joints.

Methods. Six different types of ultrasound machines were used. On each machine, the factory setting for superficial musculoskeletal scanning was used unchanged for both color and power Doppler modalities. The settings were then adjusted for increased Doppler sensitivity, and these settings were designated study settings. Eleven patients with rheumatoid arthritis (RA) with wrist involvement were scanned on the 6 machines, each with 4 settings, generating 264 Doppler images for scoring and color quantification. Doppler sensitivity was measured with a quantitative assessment of Doppler activity: color fraction. Higher color fraction indicated higher sensitivity.

Results. Power Doppler was more sensitive on half of the machines, whereas color Doppler was more sensitive on the other half, using both factory settings and study settings. There was an average increase in Doppler sensitivity, despite modality, of 78% when study settings were applied. Over the 6 machines, 2 Doppler modalities, and 2 settings, the grades for each of 7 of the patients varied between 0 and 3, while the grades for each of the other 4 patients varied between 0 and 2.

Conclusion. The effect of using different machines, Doppler modalities, and settings has a considerable influence on the quantification of inflammation

Address correspondence to Søren Torp-Pedersen, MD, Department of Diagnostics, Radiology Section, Copenhagen University Hospital at Glostrup, Nordre Ringvej 57, 2600 Glostrup, Denmark. E-mail: stor0007@regionh.dk.

Submitted for publication January 2, 2014; accepted in revised form October 28, 2014.

by ultrasound in RA patients, and this must be taken into account in multicenter studies.

In rheumatology, the increasing use of ultrasound has been driven by a need to accurately identify and suppress inflammation. To this end, the role of Doppler imaging has gained importance with respect to its ability to detect joint hyperemia, a key feature of the inflammatory process. The presence of increased Doppler signals in the joints of patients with inflammatory arthritis has been shown to predict disease persistence and progression to rheumatoid arthritis (RA) and future bone damage (1,2). The accurate assessment of Doppler imaging is therefore crucial in patients with inflammatory arthritis.

In musculoskeletal disease, Doppler studies in joints and soft tissues are usually undertaken with either power Doppler or color Doppler. Each evaluates and represents different aspects of blood flow. Power Doppler displays the energy of all moving erythrocytes at each sampling site. Color Doppler displays the direction (up or down) and the mean velocity of all moving erythrocytes at each sampling site. Following the emergence of power Doppler in the 1990s, it was considered the most sensitive Doppler modality. This was because there appeared to be a theoretical basis to explain the difference, and, in the early machines, power Doppler did appear to be more sensitive than color Doppler (3,4). The explanation was that at each sampling site, the energies of all of the different velocities are summed, generating a better signal-to-noise ratio, whereas in color Doppler, the average velocity at each sampling site is displayed. Each velocity has to rise above the noise threshold by itself. Because of the high sensitivity of power Doppler, it was even described as being angle independent. However, all Doppler modalities have to abide by the laws of physics. When the blood cells move at right angles (orthogonal) to the beam of sound, no Doppler shift is generated and therefore cannot be detected. In practice, the higher the sensitivity of the Doppler the less angle dependent it seems (3). Interestingly, new advances in technology have led to new generations of machines and the observation that color Doppler appears to be more sensitive than power Doppler in some machines (5).

Both Doppler modalities are used to monitor treatment strategy for efficacy and treatment failure in patients with rheumatic diseases, and it is of interest in rheumatology practice and in research to score the changes over time (2,6,7). For this purpose, different

scoring systems have been proposed. The most frequently used are semiquantitative scoring systems that score the Doppler information on a scale of 0–3, with increasing scores indicating increasing amounts of color in the synovium (8–10). Apart from monitoring disease activity during treatment, Doppler findings have been proposed as a predictive tool for relapse in patients in remission (11,12) and have been shown to predict erosive progression both in patients with early RA and in patients with low disease activity or remission (13–15). Using a semiquantitative scoring system has been suggested to aid in the diagnosis of RA in patients with undifferentiated arthritis by using a certain Doppler grade as a cutoff (1,16–18). With higher Doppler sensitivity, more color is displayed, and therefore higher grades are scored in comparison to a machine with lower Doppler sensitivity. Furthermore, with the ongoing development of more sensitive Doppler systems, what might be the optimal Doppler modality for slow flow (a setting that detects slow as well as fast flow, i.e., all flow—the fast flow possibly with incorrect direction and velocity) on one machine may not be the optimal modality on another. A slow flow setting is necessary to investigate synovial flow in order to ensure the detection of venous as well as arterial flow.

We therefore found it of interest to investigate the sensitivity of power Doppler and color Doppler ultrasound imaging before and after optimization on different machines and how the scoring of inflammation would be affected by the difference in Doppler sensitivity since these findings may have a major impact on treatment decisions and diagnosis.

The objectives of this study of patients with active RA were 1) to determine which Doppler modality, color Doppler or power Doppler, on different machines was more sensitive for the detection of flow in the dorsocentral aspect of the wrist joint in patients with active RA, 2) to determine if adjusting the settings provided by the manufacturers (called factory settings) for color Doppler and power Doppler increased the sensitivity, and 3) to determine if variation in the machines, Doppler modalities, and Doppler settings affected the Doppler scoring.

PATIENTS AND METHODS

Patients. Eleven patients who fulfilled the American College of Rheumatology 1987 criteria for RA were selected from 2 departments of rheumatology for the study (19). They were included based on clinical signs of active disease in the

wrist joint and subsequent Doppler examination ensuring the presence of Doppler activity in the wrist (grade 2, ranging from mild to severe within that grade). The patients were recruited in the week before the trial.

Ultrasound machines and settings. The following machines were available for the study: Siemens Acuson S2000 with probe 18L6HD, Philips IU22 with probe L17-5, Esaote MyLab 70 XVG with probe LA435, General Electric (GE) Logiq E9 with probe ML6-15, GE P5 with probe 11L, and GE P6 with probe 11L. We considered the first 4 machines to be

high end and the latter 2 to be medium range. In agreement with the ultrasound companies, the individual performances of the machines were anonymized, and in the Results, tables, and figures, they are referred to as machines 1–6, with the latter two identified as medium range. Whether color Doppler or power Doppler was the most sensitive modality was not anonymized.

On each machine, the factory setting for superficial musculoskeletal assessment was used with power Doppler and color Doppler. The settings were also adjusted for both

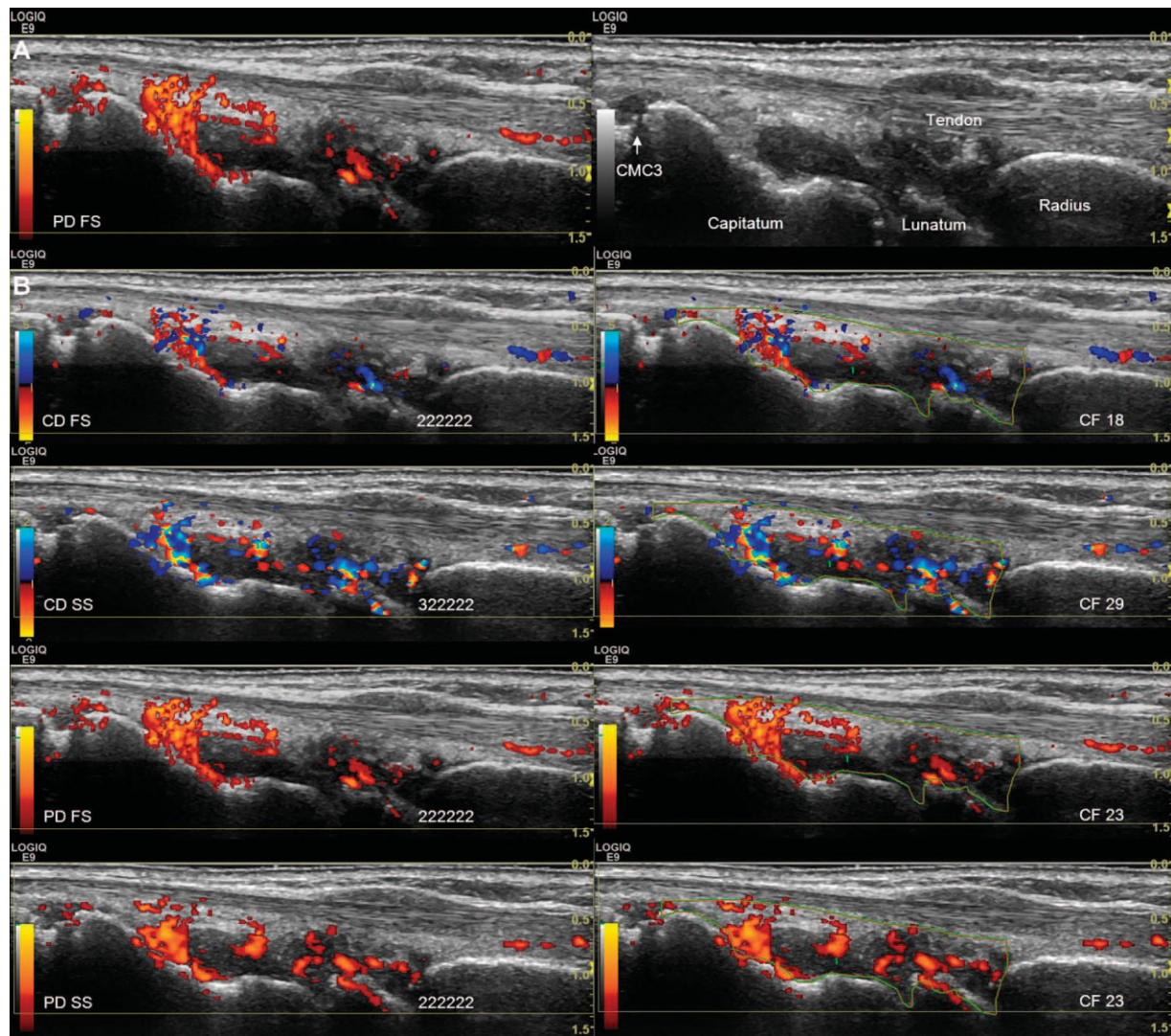


Figure 1. Example of color quantification. **A**, Left, Doppler image. Right, Corresponding grayscale image. The grayscale image is included as an aid to outline the synovium since the synovial borders may be covered by color on the Doppler image. **B**, Left, The 4 images as seen by the investigators during scoring. The scores (almost exclusively 2) are shown in the bottom right corner of each image. Right, The 4 anatomic regions of interest (ROIs). The color fraction (CF) is shown in the bottom right corner. Note that the same ROI is used in the 4 images. Images were obtained using a General Electric Logiq E9 machine. PD = power Doppler; FS = factory settings; CD = color Doppler; SS = study settings; CMC3 = third carpometacarpal joint.

modalities for increased Doppler sensitivity by decreasing the pulse repetition frequency, decreasing the wall filter, increasing the color threshold (color priority), and adjusting the Doppler gain to the level just below random noise, and these settings were saved as the study settings (5). A total of 4 settings were used: color Doppler factory setting, power Doppler factory setting, color Doppler study setting, and power Doppler study setting. All machines were set to store a 4-second video clip.

Scanning. The longitudinal dorsal plane over the wrist in line with the third finger was used. The following landmarks

were used in the image: distal radius, lunate bone, capitate bone, and extensor digitorum tendon (Figure 1). The scan plane was marked on the skin with a pen.

The patient was positioned opposite the investigator with the hand relaxed and prone on the examination couch. A generous amount of gel was applied, and the investigators were instructed to have scanning gel visible in the image to reduce transducer pressure. Respecting the scanning plane drawn on the skin and with the landmarks in the image, the Doppler was activated. Slight adjustment of the transducer position was made until the scanning plane with maximum Doppler activity

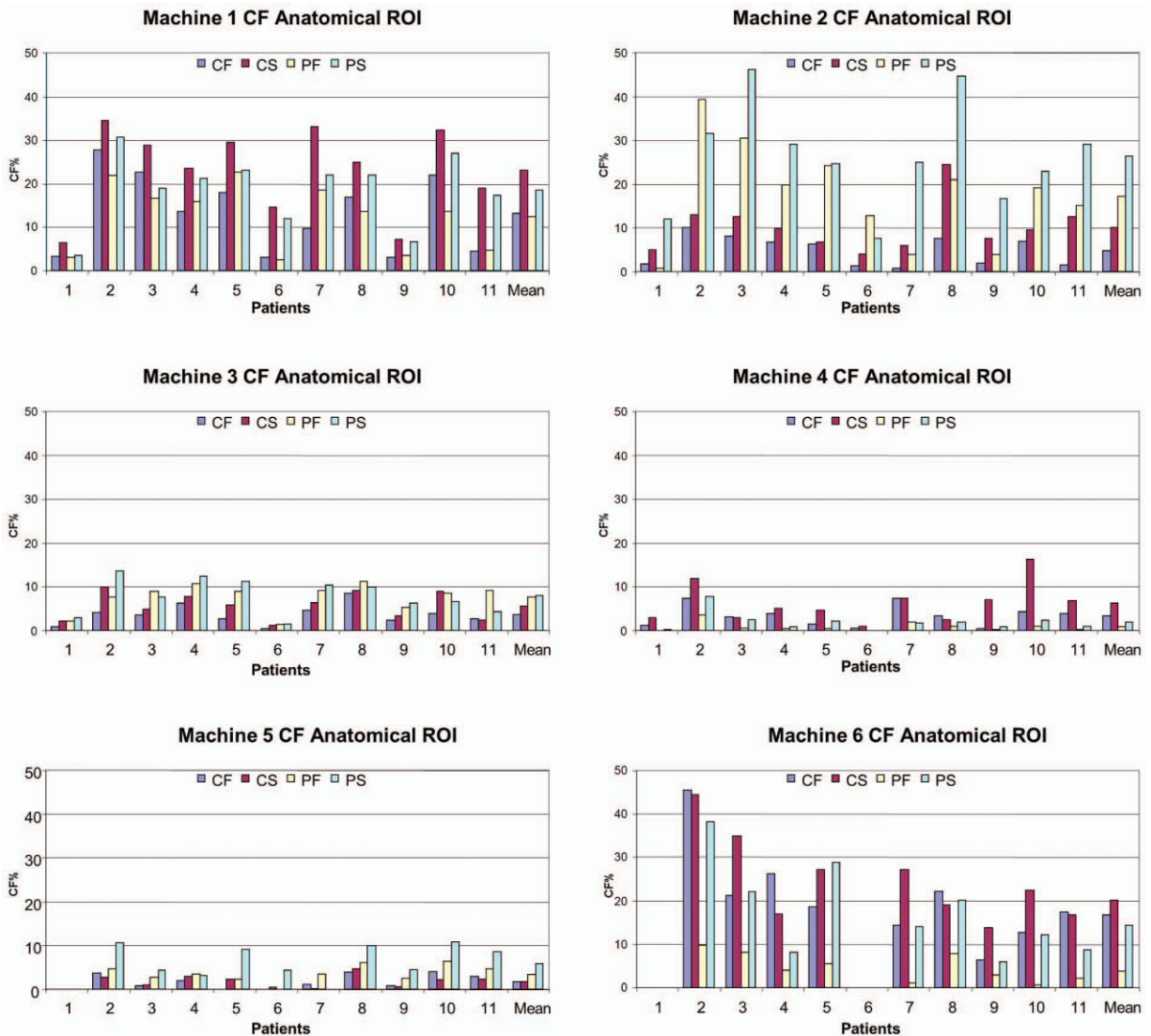


Figure 2. Results of color quantification with anatomic region of interest (ROI) Each graph shows the color fraction (CF%) with anatomic ROIs for the 11 patients who were each scanned using the 2 Doppler modalities and 2 settings. For each machine, the most sensitive Doppler modality is indicated by the highest bar. Machines 1–4 are high-end machines, and machines 5 and 6 are medium range. CF = color Doppler with factory settings; CS = color Doppler with study settings; PF = power Doppler with factory settings; PS = power Doppler with study settings.

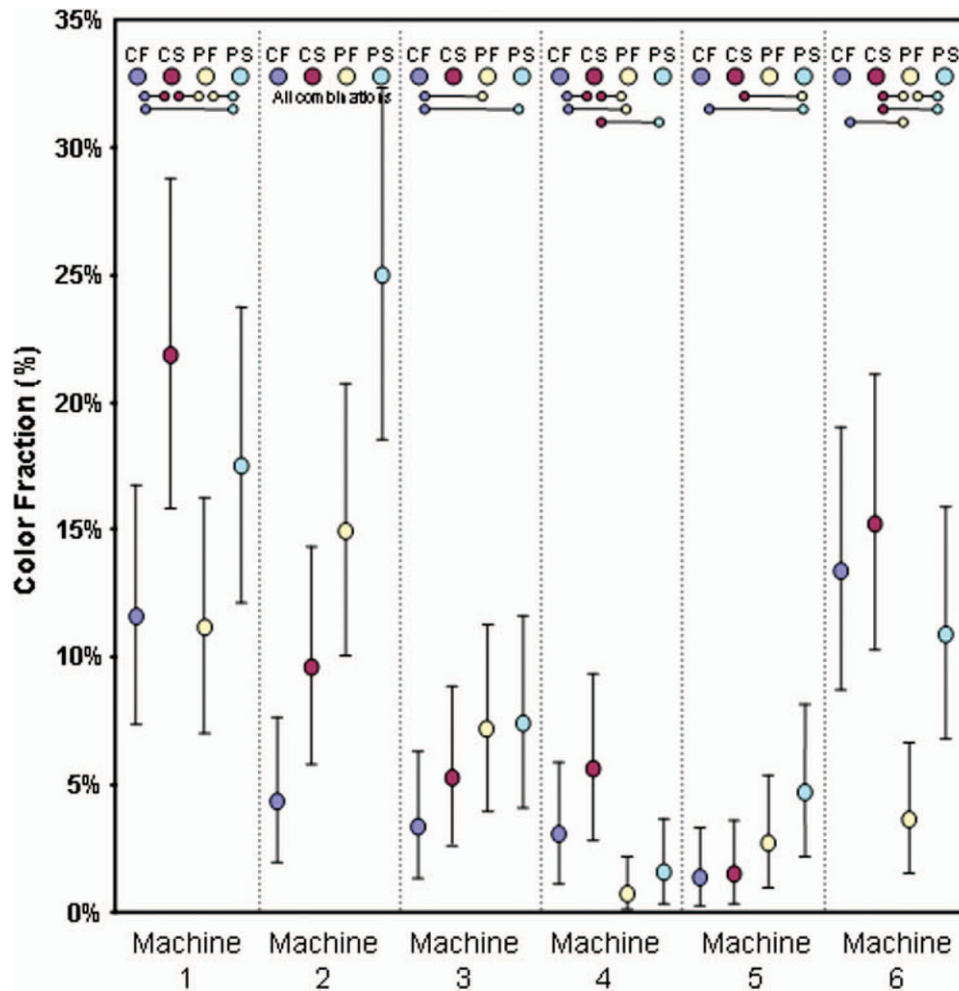


Figure 3. Average color fraction for the 6 machines with 4 settings. For each machine, the average color fraction in the 11 patients is shown for color Doppler with factory settings, color Doppler with study settings, power Doppler with factory settings, and power Doppler with study settings. Circles show the mean, and bars show the 95% confidence interval. Connected symbols at the top indicate statistically significant differences. Machines 1–4 are high-end machines, and machines 5 and 6 are medium range. See Figure 2 for definitions.

was found. The transducer was held in this position while four 4-second clips were stored with color Doppler and power Doppler under factory and study settings. This sequence was followed in all patients. Finally, a still image in the same scan position in B-mode was stored.

The 11 patients moved in a circle and were scanned on all 6 machines manned by 6 investigators (KE, MAD, AI, PB, MS, and EN) during a 90-minute rotation between 10:00 AM and noon. In this way, 264 clips were stored (11 patients on 6 machines with 4 clips each).

Image analysis. From each clip, the image with maximum Doppler activity within the synovium was selected. This image was color graded and color quantified.

Color grading. A modified Szkudlarek grading scale of 0–3 was used (8), where 0 = no color visible in the synovium, 1 = up to 3 single-vessel foci visible in the synovium, 2 = more

than grade 1 and <50% of the synovium covered by color, and 3 = >50% of the synovium covered by color.

The 264 Doppler images and 66 grayscale images were sent to 6 investigators (LT, MAD, AI, PB, MS, and EN) who returned the color grading on a form. The investigators were blinded with regard to the results of color quantification as well as the machine settings. The grayscale images were included as an aid in case the color made the borders of the synovium difficult to define. Each of the 264 images thus received 6 grades. A consensus grade was generated by selecting the most frequent grade among the 6 grades. If more than 1 grade was most frequent, the highest of these grades was selected as the consensus grade.

Color quantification. The Doppler images were imported into the program Image-Pro Analyzer version 6.3 (Media Cybernetics), and the region of interest (ROI) was

traced using an anatomic ROI. The trace followed the carpal bone surfaces (bottom) and the extensor digitorum tendon (top) between the distal tip of the radial bone (proximal) and the carpometacarpal joint (distal) (Figure 1). The ROI included the synovium as well as the loose connective tissue surrounding it. Color quantification was performed by an investigator (ATP) who was blinded with regard to the results of grading.

Interpretation. The color fraction in the anatomic ROI is used to describe Doppler sensitivity. The more flow a given machine with a given setting can display within the ROI, the higher the Doppler sensitivity. The Doppler grading is used to illustrate the clinical impact by investigating whether the 11 patients receive different grades as they move between machines, modalities, and settings.

Statistical analysis. Statistical analyses were based on mixed linear models, which is a generalization of the standard linear model used in the General Linear Model procedure. The data are permitted to exhibit correlation and nonconstant variability (i.e., repeated measures on the same participants). The null hypothesis was based on the triple interaction between ultrasound machine, setting, and the type of Doppler. The model included all of these as main effects as well as all of the combinations of interactions. The 5% significance level (2-sided) was used.

RESULTS

Doppler sensitivity. The color fractions using anatomic ROIs are shown in Figure 2. For each machine, the highest bars indicate the most sensitive Doppler modality and setting. The amount of color in the ROI is displayed for each setting (color Doppler factory settings, color Doppler study settings, power Doppler factory settings, and power Doppler study settings, in that order) for each of the 11 patients, and the mean amount of color in the ROI for the 4 settings is shown. For instance, for machine 2, the mean for color Doppler factory settings was the lowest, the mean for color Doppler study settings was higher, the mean for power Doppler factory settings was higher still, and the mean for power Doppler study settings was the highest. This indicates that color Doppler factory settings had the lowest sensitivity and power Doppler study settings had the highest sensitivity. On that machine, power Doppler was more sensitive than color Doppler, and for both modalities, the study settings were more sensitive than the factory settings.

As expected, the 11 patients displayed a varying amount of perfusion in the dorsocentral wrist, due to the selection criteria (grade 2, ranging from mild to severe within that grade). Patients 1 and 6 displayed the least perfusion, and the 2 medium-range machines (machines 5 and 6) had difficulties detecting it. Machine 4 barely detected flow in these 2 patients, and only detected it

Table 1. Effect of adjusting settings for more sensitive Doppler*

Machine and Doppler modality	Average color fraction		% improvement†
	Factory settings	Study settings	
Machine 1			
Color Doppler	13	23	75
Power Doppler	12	19	49
Machine 2			
Color Doppler	5	10	109
Power Doppler	17	26	52
Machine 3			
Color Doppler	4	6	56
Power Doppler	8	8	4
Machine 4			
Color Doppler	3	6	86
Power Doppler	1	2	134
Machine 5			
Color Doppler	2	2	0
Power Doppler	3	6	80
Machine 6			
Color Doppler	17	20	20
Power Doppler	4	14	273

* The average color fractions for the 11 patients were determined using factory settings and study settings with each Doppler modality. Machines 1–4 are high-end machines, and machines 5 and 6 are medium range.

† The average improvement was 78%.

with color Doppler. The performance of machine 6 is interesting. It did not detect flow in patients 1 and 6, indicating a low Doppler sensitivity. However, when it did detect flow, it displayed it with a high color fraction, which may incorrectly be interpreted as a high Doppler sensitivity. Machines 1–3 detected flow in all patients with all settings.

Using anatomic ROI, power Doppler was more sensitive on the Esaote MyLab 70 XVG, Siemens S2000, and GE P6, whereas color Doppler was more sensitive on the GE E9, Philips IU22, and GE P5. This was the case for both factory settings and study settings. The mean color fraction with confidence intervals for each machine as well as statistically significant differences are shown in Figure 3.

Machine settings. When study settings were compared to factory settings, there was an average increase in Doppler sensitivity of 78% (range 0–273%) (Table 1).

Doppler scoring. The consensus scores (most frequent score on each of the 264 images) are shown in Figure 4. The scores shifted left or right depending on modality and settings. The figure illustrates the clinical impact of varying Doppler sensitivities. With machines 1–3, scores of 0 and 1 were rare and almost exclusively

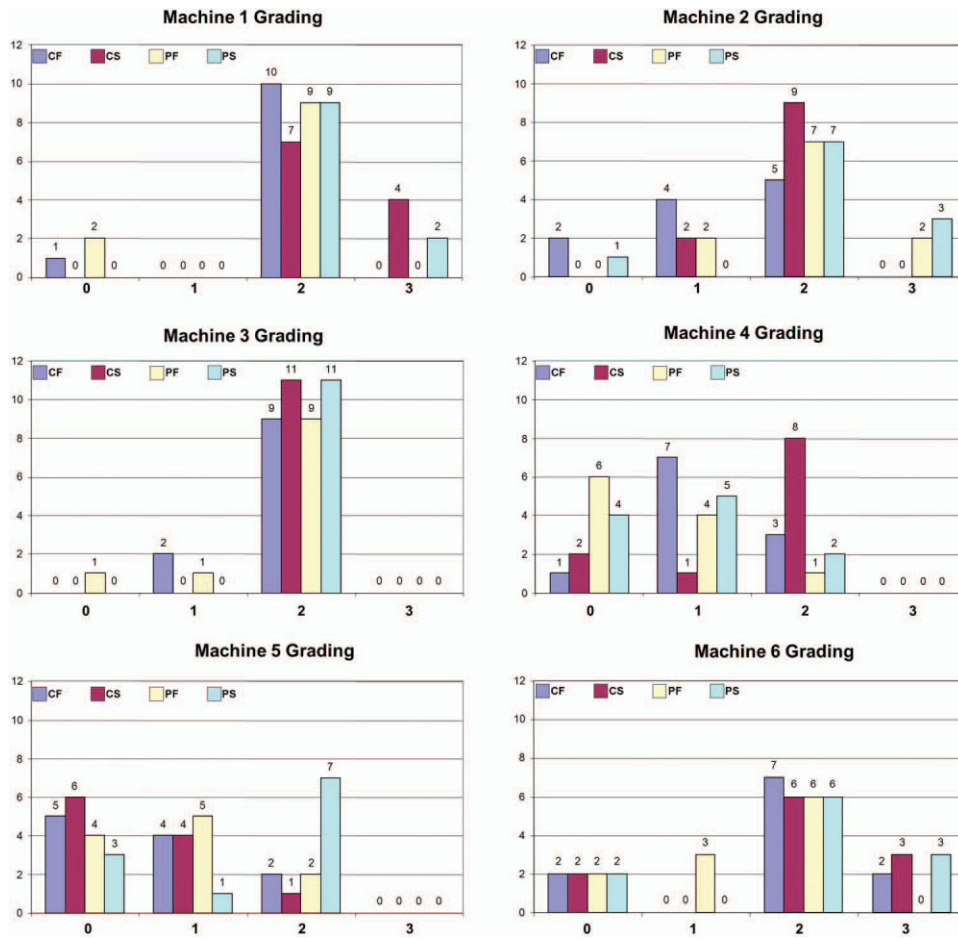


Figure 4. Results of Doppler grading on 6 machines. The distribution of grades 0–3 is shown for each setting on each machine. Machines 1–4 are high-end machines, and machines 5 and 6 are medium range. See Figure 2 for definitions.

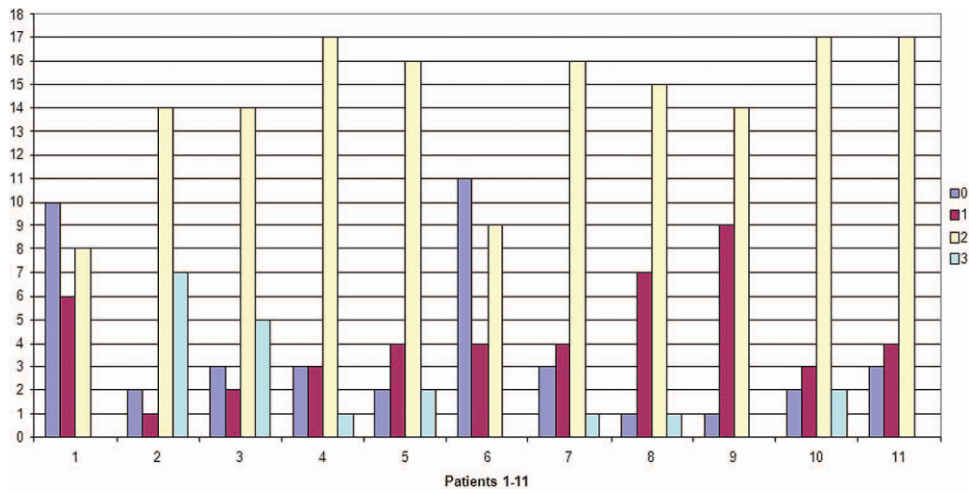


Figure 5. Distribution of Doppler grades 0–3 in the 11 patients.

occurred with factory settings. With machines 4 and 5, the low scores were common. With machine 6, a mixture of low and high scores was seen. Figure 5 shows the distribution of scores within each patient. When the 11 patients moved between machines, Doppler modalities, and settings, the grades for 7 of the patients varied between 0 and 3, and the grades for the remaining 4 patients varied between 0 and 2. Patients 1 and 6 had the least flow (Figure 2) and had grade 0 as the most common grade. In the rest of the patients, grade 2 was the most common grade.

DISCUSSION

Doppler findings have a considerable clinical impact on the management of rheumatic diseases. The accurate detection and grading of blood flow is therefore of paramount importance, since the display of more color is interpreted as more inflammation.

This is the first study to show that the use of different machines, Doppler modalities, and machine settings has a marked influence on the apparent degree of inflammation. Indeed, the grades for 7 of the patients moved between 0 and 3, and the grades for the remaining 4 patients moved between 0 and 2. This means that, depending on the methodology used, the same patient may be assessed as having no inflammation or having up to considerable inflammation using a modification of the 4-point scale originally described by Szkudlarek (8). It underlines the importance of knowing a machine's Doppler performance. This study also shows that it is worthwhile to adjust the Doppler parameters. The average improvement in Doppler sensitivity of 78% is substantial, and the improvement was even as high as 273% for one machine. The factory settings are not at the highest sensitivity per se, and users of the machines must know how to optimize the Doppler to obtain optimal flow information (5). These findings indicate that the suggested cutoff values for Doppler activity for diagnosing RA (1,16,17) may not be valid in all institutions, since the grades will depend on Doppler sensitivity. Likewise, Doppler findings in patients in remission may not be seen if the Doppler sensitivity is low, and the predictive value for possible flare may not be found, which could have a major impact on determining whether to reduce drug therapy (11,12,15).

In the present study, half of the machines showed that color Doppler was more sensitive than power Doppler, and the other half showed that power Doppler was more sensitive than color Doppler. This was the case

with the 4 high-end machines as well as the 2 medium-range machines and indicates that power Doppler should not be chosen as the default mode. Instead, which mode is more sensitive needs to be established on the individual machine. In order to understand why color Doppler might be more sensitive, we wrote to the engineers at 4 leading companies for an explanation. We received responses from 3, and each gave a similar response, which is summarized below (MacQuarrie J [Philips], Maccio M [Esaote], Kristoffersen K [General Electric]: personal communication).

It is a misconception that power Doppler is inherently more sensitive than color Doppler. Both have basically the same detection process and rely on the same physics (e.g., the Doppler equation) and are thus similarly affected by factors such as Doppler angle, ensemble length, burst length, etc. Both also use essentially the same signal processing (e.g., autocorrelation, wall filtering, segmentation, etc.).

The difference is which parameter extracted from the processing is displayed on the image—power in one case and velocity (and variance) in the other. However, both techniques allow tradeoffs to be made between sensitivity and other image properties. The Doppler sensitivity depends on many variables, each of which may be adjusted differently in different models and software programs. The end result is that the sensitivity of a given Doppler configuration cannot be predicted from the design but has to be determined in practice.

For one company it is even a general rule to set the factory parameters so that the power Doppler mode is more sensitive than the color Doppler, since a power Doppler mode that is less sensitive than the color Doppler mode would be not useful. So either color Doppler or power Doppler may be the more sensitive modality, either by chance or due to a deliberate choice by the manufacturer. These explanations, taken together with our findings, underline the fact that the most sensitive Doppler modality must be determined in practice and is not given by default.

Several limitations of the study need to be considered. The patients were selected from 2 departments of rheumatology, and in order to be included in the study, they had to have visible hyperemia (at least grade 2) in the dorsocentral wrist, as determined using a GE Logiq E9 or Logiq9 ultrasound machine, which were the machines routinely in use in the 2 departments of rheumatology. In order to rank the 2 Doppler modalities, using 2 settings for each in 6 different machines, in terms of ability to display flow, we used the perfusion in

the dorsocentral portion of the wrist in 11 patients with active inflammatory RA. We do not know what the actual perfusion was that day in those patients but find it likely that the perfusion was unaltered during the 90 minutes of the trial (20). Also, the findings of this study were all obtained using the dorsal aspects of the central wrist, and it is possible that studies in other anatomic regions may show different results.

We are aware that the transducer pressure may affect the amount of perfusion displayed. Therefore, all machines were manned by trained investigators who knew the importance of a very light transducer pressure. We chose to measure the amount of color within an ROI defined by the extensor digitorum tendon and bony landmarks. This ROI is easily defined in all patients and on any machine because it is not affected by difficulties in defining the synovial outline and is therefore subject to less variation (21). Also, the study was a sensitivity study, not a reliability study on image acquisition. We used the amount of color displayed in the ROI as a measure of perfusion and thereby as a measure of sensitivity. This is, however, not entirely accurate. The amount of color in the ROI is also affected by blooming, slice thickness, and spatial resolution of the Doppler.

All machines have blooming artifacts with both modalities and both settings. We did not try to avoid this because minimizing the blooming artifact minimizes the Doppler sensitivity, and flow information is lost. When we adjusted the settings to create study settings, we increased the sensitivity to the level where movement artifacts became difficult to avoid, and as a consequence of the high sensitivity, blooming was maximized. Some machines had more blooming than others, and on some machines, there was a different level of blooming between power Doppler and color Doppler. Nevertheless, we accepted the presence of blooming as a systematic error since it is flow generated and thus not entirely artifactual.

The height of the bars in Figure 2, which show the color fraction, cannot be used to determine if one machine is more sensitive than another. A machine with a sensitive Doppler, with a high Doppler resolution (small color pixels), with a thin image plane, and with a low degree of blooming may display less color in the ROI than a machine with a less sensitive Doppler, with low color resolution (larger color pixels), with a thick image plane (more vessels being sampled), and with a higher degree of blooming. If we had wanted to rank the machines in terms of Doppler sensitivity, we would have needed a known perfusion, a Doppler phantom. A

phantom mimicking the flow we wished to detect in soft tissue inflammation, slow flow at the arteriole level does not exist at present. Even so, we believe our results are important and relevant, as they demonstrate significant machine-, modality-, and setting-dependent variations in Doppler sensitivity and display that need to be identified by users, especially if findings from different institutions are pooled or compared.

This study shows that there is a wide variation in sensitivity to display perfusion when machines, Doppler modalities, and settings are compared. The variation is considerable and may have clinical impact since the same patient may be graded from 0 to 3 depending on these variables. It demonstrates that power Doppler is not necessarily the most sensitive modality and should be chosen as the default mode only after evaluating the machine. Furthermore, in order to obtain the highest Doppler sensitivity, it is necessary to adjust the factory settings. This study emphasizes the need to use comparable machines in multicenter studies.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. S. Torp-Pedersen had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. S. Torp-Pedersen, Christensen, Szkudlarek, Ellegaard, D'Agostino, Iagnocco, Naredo, Balint, Wakefield, A. Torp-Pedersen, Terslev.

Acquisition of data. S. Torp-Pedersen, Christensen, Szkudlarek, Ellegaard, D'Agostino, Iagnocco, Naredo, Balint, Wakefield, A. Torp-Pedersen, Terslev.

Analysis and interpretation of data. S. Torp-Pedersen, Christensen, Szkudlarek, Ellegaard, D'Agostino, Iagnocco, Naredo, Balint, Wakefield, A. Torp-Pedersen, Terslev.

REFERENCES

1. Freeston JE, Wakefield RJ, Conaghan PG, Hensor EM, Stewart SP, Emery P. A diagnostic algorithm for persistence of very early inflammatory arthritis: the utility of power Doppler ultrasound when added to conventional assessment tools. *Ann Rheum Dis* 2010;69:417-9.
2. Taylor PC, Steuer A, Gruber J, Cosgrove DO, Blomley MJ, Marsters PA, et al. Comparison of ultrasonographic assessment of synovitis and joint vascularity with radiographic evaluation in a randomized, placebo-controlled study of infliximab therapy in early rheumatoid arthritis. *Arthritis Rheum* 2004;50:1107-16.
3. Bude RO, Rubin JM, Adler RS. Power versus conventional color Doppler sonography: comparison in the depiction of normal intrarenal vasculature. *Radiology* 1994;192:777-80.
4. Rubin JM, Bude RO, Carson PL, Bree RL, Adler RS. Power Doppler US: a potentially useful alternative to mean frequency-based color Doppler US. *Radiology* 1994;190:853-6.
5. Torp-Pedersen ST, Terslev L. Settings and artefacts relevant in

- colour/power Doppler ultrasound in rheumatology. *Ann Rheum Dis* 2008;67:143–9.
6. Terslev L, Torp-Pedersen S, Qvistgaard E, Danneskiold-Samsøe B, Bliddal H. Estimation of inflammation by Doppler ultrasound: quantitative changes after intra-articular treatment in rheumatoid arthritis. *Ann Rheum Dis* 2003;62:1049–53.
 7. Naredo E, Moller I, Cruz A, Carmona L, Garrido J. Power Doppler ultrasonographic monitoring of response to anti-tumor necrosis factor therapy in patients with rheumatoid arthritis. *Arthritis Rheum* 2008;58:2248–56.
 8. Szkudlarek M, Court-Payen, Jacobsen S, Klarlund M, Thomsen HS, Ostergaard M. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. *Arthritis Rheum* 2003;48:955–62.
 9. Stone M, Bergin D, Whelan B, Maher M, Murray J, McCarthy C. Power Doppler ultrasound assessment of rheumatoid hand synovitis. *J Rheumatol* 2001;28:1979–82.
 10. Filippucci E, Farina A, Carotti M, Salaffi F, Grassi W. Grey scale and power Doppler sonographic changes induced by intra-articular steroid injection treatment. *Ann Rheum Dis* 2004;63:740–3.
 11. Scire CA, Montecucco C, Codullo V, Epis O, Todoerti M, Caporali R. Ultrasonographic evaluation of joint involvement in early rheumatoid arthritis in clinical remission: power Doppler signal predicts short-term relapse. *Rheumatology (Oxford)* 2009;48:1092–7.
 12. Foltz V, Gandjbakhch F, Etchepare F, Rosenberg C, Tanguy ML, Rozenberg S, et al. Power Doppler ultrasound, but not low-field magnetic resonance imaging, predicts relapse and radiographic disease progression in rheumatoid arthritis patients with low levels of disease activity. *Arthritis Rheum* 2012;64:67–76.
 13. Lillegraven S, Boyesen P, Hammer HB, Ostergaard M, Uhlig T, Sesseng S, et al. Tenosynovitis of the extensor carpi ulnaris tendon predicts erosive progression in early rheumatoid arthritis. *Ann Rheum Dis* 2011;70:2049–50.
 14. Naredo E, Collado P, Cruz A, Palop MJ, Cabero F, Richi P, et al. Longitudinal power Doppler ultrasonographic assessment of joint inflammatory activity in early rheumatoid arthritis: predictive value in disease activity and radiologic progression. *Arthritis Rheum* 2007;57:116–24.
 15. Brown AK, Conaghan PG, Karim Z, Quinn MA, Ikeda K, Peterfy CG, et al. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. *Arthritis Rheum* 2008;58:2958–67.
 16. Filer A, de Pablo P, Allen G, Nightingale P, Jordan A, Jobanputra P, et al. Utility of ultrasound joint counts in the prediction of rheumatoid arthritis in patients with very early synovitis. *Ann Rheum Dis* 2011;70:500–7.
 17. Nakagomi D, Ikeda K, Okubo A, Iwamoto T, Sanayama Y, Takahashi K, et al. Ultrasound can improve the accuracy of the 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis to predict the requirement for methotrexate treatment. *Arthritis Rheum* 2013;65:890–8.
 18. Salaffi F, Ciapetti A, Gasparini S, Carotti M, Filippucci E, Grassi W. A clinical prediction rule combining routine assessment and power Doppler ultrasonography for predicting progression to rheumatoid arthritis from early-onset undifferentiated arthritis. *Clin Exp Rheumatol* 2010;28:686–94.
 19. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
 20. Semerano L, Gutierrez M, Falgarone G, Filippucci E, Guillot X, Boissier MC, et al. Diurnal variation of power Doppler in metacarpophalangeal joints of patients with rheumatoid arthritis: a preliminary study. *Ann Rheum Dis* 2011;70:1699–700.
 21. Ellegaard K, Torp-Pedersen S, Lund H, Henriksen M, Terslev L, Jensen PS, et al. Quantification of colour Doppler activity in the wrist in patients with rheumatoid arthritis—the reliability of different methods for image selection and evaluation. *Ultraschall Med* 2008;29:393–8.