Ultrasound in Rheumatology

Fundamental background and clinical applications

DAVID F. TEN CATE

Ultrasound in Rheumatology

Fundamental background and clinical applications

Echografie in reumatologie

Fundamentele achtergronden en klinische toepassingen

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

prof.dr. H.A.P. Pols (rector magnificus)

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

8 april 2015 om 09.30 uur

David Folkert ten Cate
geboren te Rotterdam

ERASMUS UNIVERSITEIT ROTTERDAM

PROMOTIECOMMISSIE

Promotor(en):
    Prof. Dr. J.M.W. Hazes

Overige leden:
    Prof. Dr. G.J. Kleinrensink
    Prof. Dr. Ir. N. de Jong
    Prof. Dr. R. Landewé

Copromotor(en):
    Dr. J.J. Luime
    Dr. J.W.G. Jacobs

Voor Miek, Luuk en Gijs

Table of contents

Assessment of disease activity in patients with rheumatoid arthritis using OST measurements, a non-invasive operator-independent imaging technique. Submitted

CHAPTER 1

General introduction
and outline of thesis

Ultrasonography is an imaging modality that is used increasingly in daily practice for the investigation and management of rheumatoid arthritis (RA). This introduction provides an overview on the rationale for using US in RA by discussing relevant events and developments in the history of RA. This is followed by a short section on the history of medical US. Subsequently, the section on the technique of US will describe relevant concepts of US for this thesis. This general introduction closes with the outline of the thesis and the aims.

RATIONALE TO USE US IN RA

RA is a common disease with a prevalence among adults of 0.5-1% worldwide; the prevalence varies geographically however (1). In the Netherlands the prevalence is 0.5% for men and 0.9% for women. In 2011 in the Netherlands, the incidence was 54 per 100,000 for men and 80 per 100,000 for women; the mortality rate was 0.34 per 100,000 men and 0.93 per 100,000 women (2).

Diseases involving inflamed joints and resulting in bone erosions, such as RA, are ancient. There is paleopathological data that indicates that erosive polyarthritis has been present as early as 3000 BCE (3). The distinction between the respective rheumatic diseases was absent for very long, however. The first to explicitly describe an affliction we may call RA nowadays is Lopez de Hinojosos. He described in 1578 a chronic arthritis always attacking the same joints, without gouty tophi, and being associated with muscular atrophy (4). This disease was not yet called RA. Garrod introduced the name 'rheumatoid arthritis' in 1859, and made a distinction between this disease and osteoarthritis and gout (5). Nevertheless, until the discovery of the association of the HLA-B27 antigen with ankylosing spondylitis in 1973 (6, 7) there was discussion whether RA and ankylosing spondylitis were one entity or not; the debate whether RA and psoriatic arthritis are two different diseases lasted even longer.

Even today, it is felt that RA is a heterogeneous disease and even best considered a clinical syndrome, rather than one distinct disease (8). We do not know the cause of RA, which means there is no clear diagnostic test. The accepted way to define RA is therefore by means of classification criteria (9). These are often used as diagnostic criteria, although not designed for that purpose.

Early RA often is difficult to diagnose, because of lack of typical signs and symptoms in the early phase, difficult clinical assessment of early synovitis and limited diagnostic value of laboratory tests. In established RA, the assessment of disease activity is difficult as well. For instance, the disease activity score assessing 28 joints (DAS28) frequently used to score RA disease activity in individual RA-patients has been validated

for evaluations at the group level, i.e. for measuring effects in clinical trials. It has not been validated for evaluations at the individual patient level, i.e. for use in clinical practice, and has several drawbacks for this use. Specific examples of these drawbacks are underestimating the disease activity, e.g. by an insensitive clinical exam, or overestimating the disease activity, e.g. in concomitant pain syndrome or osteoarthritis (10, 11). Especially for diagnosing remission DAS28 may be not valid enough (11).

In need of a more objective tool for diagnosing early RA and assessing disease activity, the use of US has been introduced in clinical practice. Its potential value has been recognised in the publication of the 2010 classification criteria for RA (9) and is corroborated by research in which US predicted progression to RA in undifferentiated arthritis (UA) (12, 13). Also in RA patients in remission US had added value to conventional measures in predicting radiographical progression of joint damage and occurrence of disease flare (14-17). The added value of US as a monitoring tool in treatment strategies in established RA patients has not been investigated in depth yet, however.

HISTORICAL BACKGROUND OF ULTRASOUND

To appreciate the road that was travelled to come to the medical US devices we have today I proceed with a short section on the history of US. A major milestone in the history of medical ultrasound is the discovery of piezoelectricity. The effect was described in 1880 by the brothers Pierre and Jacques Curie, as a property of certain materials to generate an electrical charge as a response to a mechanical stimulus. (Figure 1) Soon after, it was deduced by Lippmann that the inverse piezoelectric effect also exists, i.e. that an electrical charge results in a mechanical distortion, which was immediately confirmed experimentally by the brothers Curie (18). Currently, piezoelectric material is used in a range of applications, from electric cigarette lighters (accumulation of
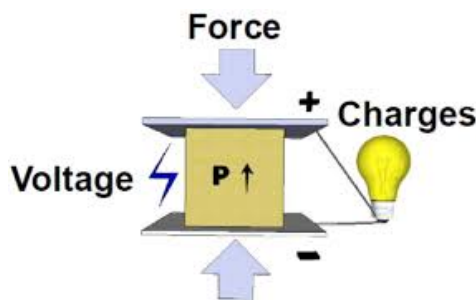
Figure 1: The piezoelectric effect: When applying a mechanical stimulus (Force) to a piezoelectric material (P) a voltage is created.

1

electrical charge when the crystal is struck) to loudspeakers (electrical charge creates mechanical movement).

In the beginning of the 20th century piezoelectrical material was used to develop a 'sound navigation and ranging' device (SONAR) by Paul Langévin to detect German U-boats. This device, also called a 'hydrophone' gave of a high pitched sound. Based on the time it took for the echo of the sound to return, the distance to a foreign object, e.g. a German U-boat, could be calculated. During the experiments which led to Langévin's SONAR it was already noted that the potent sound beam could kill fish (19), an observation which would be a preview for the emergence of therapeutic medical ultrasound, which is described below.

A parallel development, which would also prove to be essential for the emergence of medical ultrasound was the development of metal flaw detection by ultrasound. (Figure 2) Sergei Sokolov suggested in 1928 that ultrasound could be a feasible method for the detection of flaws in metal. The principles underlying this idea are the same as those of the medical ultrasound still employed today. The idea was to send ultrasonic vibrations into a metal object. If the material was without cracks a signal would return from the outer surface of the material and one from the back wall. In the case of a crack, however, an additional signal pulse would return. (Figure 3) This was a very welcome technique in the interbellum, a turbulent period in which the integrity of ships, battle tanks and bombshells was of high importance. The second World War accelerated these developments even further (20).

Combining these two parallel developments led to the development of medical ultrasound. The observations by Langévin, regarding the potential destructive effect of ultrasound, eventually led to the treatment of, e.g. Parkinson patients, using high intensity focused ultrasound as a neurosurgical tool (21). Ultrasonographic energy
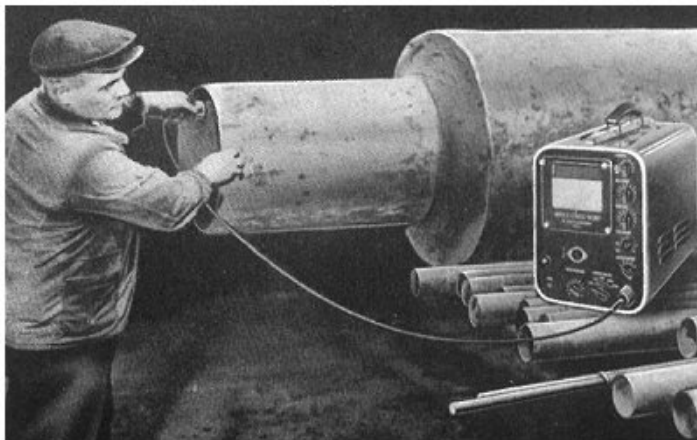
Figure 2: A worker evaluating a metal object for flaws using US
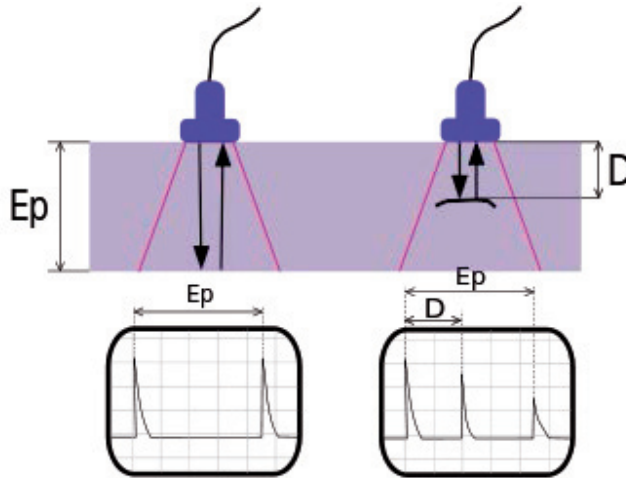
Figure 3: On the left hand side metal without cracks, displaying two signals on the screen. On the right hand side a crack is present creating an additional signal.

at a lower level was used therapeutically for a range of conditions, amongst which rheumatological conditions (22). The developments in the metal industry regarding the evaluation of the integrity of metal objects using US led to the diagnostic medical US. After an attempt to use US diagnostically in the 40's by Dussik (23) there was no turning back for diagnostic medical US from the beginning of the 50's. Wild and Reid described in 1952 the construction of an ultrasound machine able to determine structures in superficial tissues (24). Another one of the pioneers was Ian Donald, who in 1955 borrowed a metal flaw detection device, slightly modified it, and proved that one could differentiate between different abdominal masses (25). From here on US disseminated through almost all medical specialties. US in rheumatology remained scarce however for decades (26).

FUNDAMENTAL BACKGROUND OF ULTRASOUND

Before focusing on the position of US in rheumatology it is necessary to introduce some fundamental concepts of US, that will be used or will be investigated in this thesis. It is also intended to be a very concise introduction of the fundamental concepts of US for health professionals who use US in daily clinical practice.

US is an imaging technique that uses sound waves to construct an image. The most beautiful example of US is visualizing new life, whilst still in utero (Figure 4), but it is also possible to use US for the musculoskeletal tract. (Figures 5 and 6) It is called

Figure 4: A perfect boy, Luuk.

ultrasound, because it is above the range of the human hearing (20-20,000Hz). For musculoskeletal (MSK) US scanning we generally use frequencies of 5 MHz or more, i.e. more than 5,000,000 Hz. Characteristics of sound waves, other than frequency (f (Hertz(Hz)), relevant for this thesis are the amplitude (A (Pascal)) and the wavelength (λ (meter)). The wavelength is linked to the frequency and the velocity of the sound wave (c (meter/second)) in the respective media according to the following relation: λ = c / f. (Figure 7)
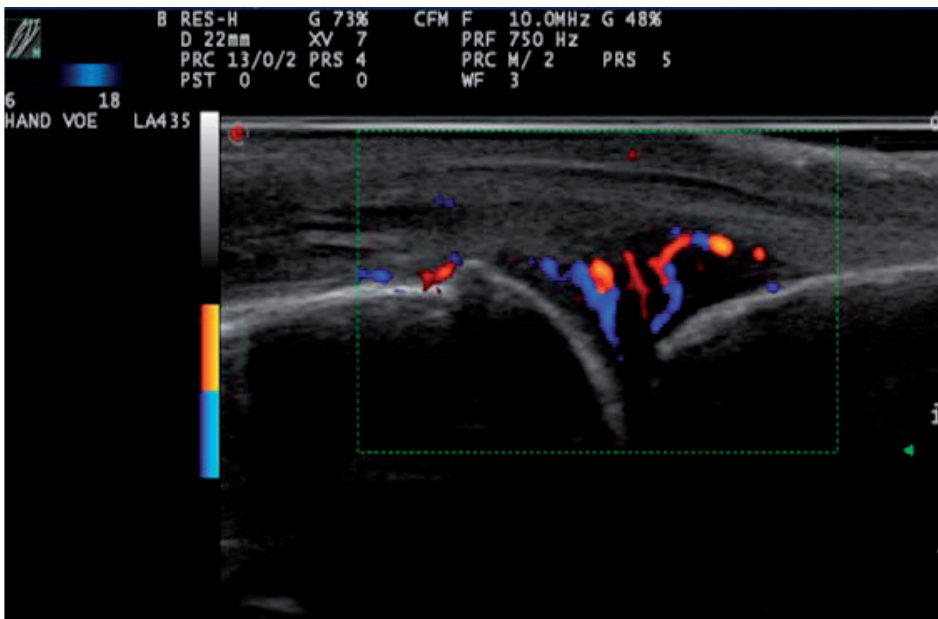
Figure 5: MCP 2 joint of a newly diagnosed male RA patient. In the joint presence of Doppler signal. The almost horizontal white line on the right hand side of the image is the proximal phalanx; from the middle of the image to the left is the caput of the metacarpal bone.
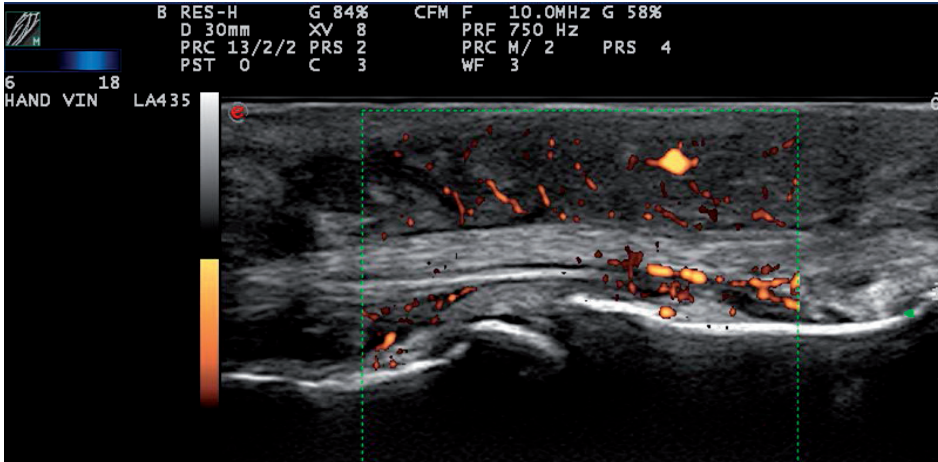
Figure 6: Flexor tendon over the second PIP joint with extensive tenosynovitis and Doppler signal
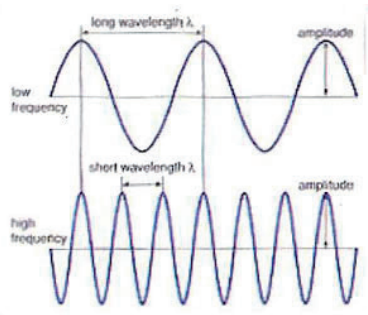
Figure 7: A high and a low frequency wave and hence different wavelengths. Amplitude is independent of these.

The sound waves needed in US scanning (Figure 8 for scanner) are produced by piezoelectric material fitted inside the transducer. This piezoelectric material changes shape in response to a high frequency alternating current. This changing shape creates a pressure wave in the medium, which is the US wave with the frequency of the applied alternating current. This sound wave is scattered (partially reflected in all directions) when it hits an inhomogenicity and partly bounced back to the transducer where it hits the piezoelectric material and is converted to an alternating current. So then the reflected sound wave acts as a mechanical stimulus on the piezoelectric material, causing an electric voltage over the material, i.e. the piezoelectric effect. This piezoelectric material is put in the transducer as several, separate crystals. These separate elements can be activated simultaneously creating a parallel, synchronous wave front. However, by activating these elements at short time intervals from each other it is possible to steer the beam and even create a focal point (Figure 9), increasing image quality. The computer can then make an image out of the electrical signals

Figure 8: A modern US scanner. At the right hand side of the horizontal control panel there are 3 transducers. Directly underneath the screen is the vertical control panel. This modern scanner is mobile enough to move around.

on the basis of the amplitude of the signal and the time it took for the echo to return to the transducer. By steering the sound beam over the tissue, a complete image is created line by line.

The sound waves travel well through homogeneous media, or from one medium to another if the acoustical impedances, which can be seen as the resistance to a sound wave, and speed of sound in the media are comparable. Fortunately, all soft tissues in humans have similar acoustic impedance and speed of sound; they are all similar to water (27). However, air and bone have very different impedance and speed of sound as compared to water. The difference in scattering properties of each respective tissue makes it possible for the US machine to construct an image with hyperechoic regions (much scattering), hypoechoic regions (less scattering) and anechoic regions (no scattering). It is very hard for the ultrasound to pass tissue-bone or tissue-air interfaces: most sound energy is reflected at the interface. When scanning without contact gel, there will be some air between probe and skin, and virtually all sound
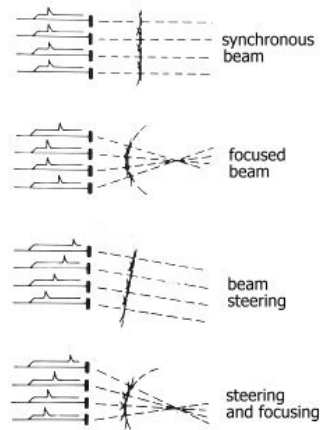
Figure 9: Beamsteering and focusing of US

waves will be reflected to the transducer immediately. This explains why you need the watery contact gel between probe and skin.

Attenuation is a characteristic of sound in media that needs to be discussed as well. As a sound wave travels through tissue it loses some amplitude per centimetre, which is attenuation. This is directly proportional to the frequency of the sound wave. (Figure 10) High frequency sound waves attenuate faster, thus less sound is reflected to the transducer; hence only superficial structures can be evaluated. However, the advantage of high frequency sound waves is that they have smaller wavelengths and thus give a higher resolution. There are two different types of resolution: axial and lateral resolution. The axial resolution describes how well two objects, lying near each other parallel to the ultrasound beam, can be distinguished one from the other. Lateral resolution describes how well two objects, lying near each other, but perpendicular to the sound beam, can be distinguished from each other. So when scanning, it is very important to decide beforehand which structure you will be scanning and how much detail you need to answer your clinical question. In other words, decide beforehand with which frequency you will be performing the US examination, taking into account the expected attenuation and necessary resolution.

With all of the above one can make a greyscale (GS) image, producing different shades of grey depending on the scattering properties of the tissue. However, one other very important modality on the US machine is the Doppler modality. There are two types of Doppler imaging relevant for rheumatological US: colour Doppler (CD) and power Doppler (PD). When sound hits a moving object, the frequency of the returning wave is changed. If the object travels away from the transducer, the frequency will be lower and if the object travels towards the transducer, it is higher. This is called the Doppler effect. This Doppler shift gives information on the velocity
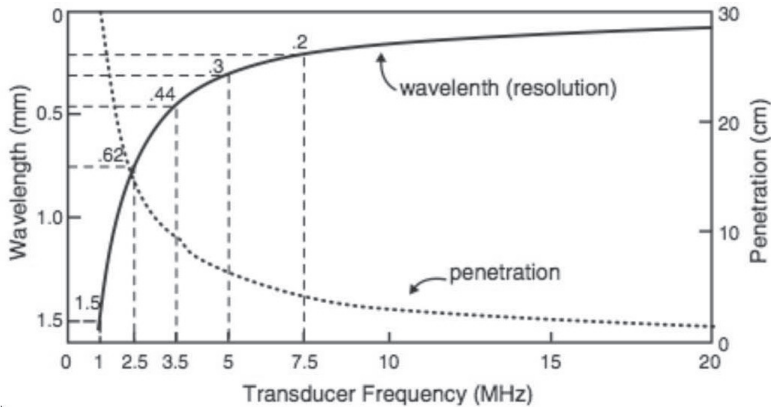
Figure 10: Higher frequency gives a better resolution, but at the cost of less penetration due to attenuation.

and direction of the particles. In practice, the Doppler shift is estimated by sending a number of pulses along the same line and estimating the phase shift between pulses. This Doppler signal is displayed as a colour overlay over part of the GS image. Since, ideally, only blood moves when performing a (MSK) US exam, blood vessels are colored, whereas surrounding tissue stays grey (no Doppler shift, so no signal). In colour Doppler, the blood flow towards the transducer is generally shown in red hues, away from the transducer in blue hues.

The other Doppler modality relevant to us is PD. In this technique once a Doppler shift is detected all frequency components, and hence velocity and direction information, are removed through integration of the power density (28). This integration of Doppler power makes PD more sensitive to low flows than CD. Also, since the background noise is displayed as a uniform, small band of colour, it can be filtered out quite easily. (Figure 11) This way, the gain of the power Doppler can be increased, making PD yet even more sensitive to low flows (27). This is why it is used more often than CD in rheumatology. We try to detect the very low flows in the neovascularised tissue present in (subclinical) joint and tendon inflammation. Also, when trying to detect this inflammation we are not interested in velocity or direction of the flow, information CD would give us, we are only interested in presence or absence.

One of the most important settings when using PD, deserving separate mentioning, is the pulse repetition frequency (PRF). PRF is a measure describing the number of pulses per second (i.e. a frequency) sent by the transducer. It is important because it determines the lowest flow velocities that can still be detected by the US machine. With a lower PRF lower possible flow velocities can be detected. Since we are only interested in low flows in rheumatology, because (subclinical) inflammation has low flows, we need to set the PRF to a low value. However, a drawback of a low PRF is more
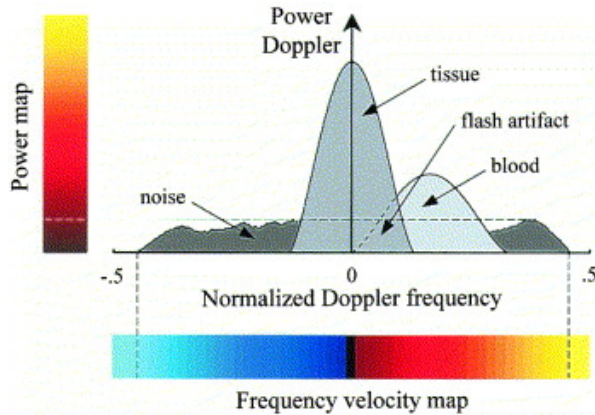
Figure 11: The advantage of power Doppler over color Doppler. In PD mode it is easier to filter out the noise than in CD mode.

motion artefacts. Because with a lower PRF the machine will become more sensitive to movement, but it will also become more sensitive to involuntary movements of the hand that is being scanned or the hand holding the transducer. For this reason, it is usually not possible in daily practice to set the PRF to the lowest value. From figure 11 we see that the Doppler power of tissue is much larger than that of blood. This is not a problem since it can be filtered out by the wall filter of the machine. The wall filter is a setting of the machine designed to filter out Doppler signal derived from tissue, especially walls of vessels. The wall filter is directly linked to PRF; a decrease in PRF allows lower levels of the wall filter (29). However, when decreasing the PRF, wall filters may be set to such a low level the motion artefacts may not be filtered out anymore, resulting in a power signal displayed on the screen. In theory this can be mistaken for true signal. However, the signal has a typical appearance, so the examiner will not mistake it for true flow (30). Nonetheless, it may interfere with the ultrasound exam as it could mask low flows occurring in inflamed tissue. So executing a proper PD exam is about selecting the best combination of settings of gain (as high as possible), wall filter (as low as possible) and PRF (as low as possible) that maximizes sensitivity to low flows and minimizes background noise or motion artefacts (31, 32).

ULTRASOUND IN RHEUMATOLOGY

Many papers appeared from 1950 onwards on the *therapeutic* use of ultrasound in rheumatology for soft tissue lesions, but it took until 1972 for the first paper to be published related to rheumatology on the *diagnostic* use of ultrasound (33). US was

1

primarily used for detection of popliteal cysts, and at the end of the 70's also for supra-patellar effusions (34). From the beginning of the 1980's there was a constant increase in the publications of the use of US in rheumatology, but the use in clinical practice remained rather limited (26). For more than a decade ultrasound in rheumatology remained limited to large joints and bursae, probably because in those years, frequencies of ultrasound transducers ranged around 5 to a maximum of 7.5 MHz. These relatively low frequencies did not allow detailed imaging of small joints, often the first affected in arthritis and RA. Not until the beginning of the 90's, higher frequency transducers were introduced (35), which was around the same time, serendipitously, when the introduction of power Doppler took place (36).

These developments created the conditions for an accelerated dissemination of US in rheumatology, resulting in a steep increase in published papers on the topic in the second half of the 90's. From the start of the current millennium this increase continued. The position of ultrasound in RA took shape with Wakefield showing the higher sensitivity of ultrasound for detecting bone erosions and soft tissue lesions, compared to conventional radiography and clinical examination, respectively (37, 38). Ultrasonographic findings in the small hand joints were validated using an magnetic resonance imaging (MRI) scanner (39), and the first steps were taken to use US for monitoring treatment response (40). After these endeavours, determining the exact place of US in rheumatology has become more centralized under the wings of OMER-ACT (Outcome measures in rheumatology), an 'independent initiative of international health professionals interested in outcome measures in rheumatology' (41). The OMERACT has done great work in defining what ultrasonographic inflammation is (42) and, since US is an operator dependent technique, in increasing the reliability of the technique (43).

ISSUES IN RESEARCH IN ULTRASOUND IN RHEUMATOLOGY

Despite the progress ultrasonographic research has made in rheumatology, still important issues remain. For example, regarding validity, data on histology and anatomical dissection as reference to validate greyscale ultrasound (GSUS) and power Doppler ultrasound (PDUS) is still scarce. Common definitions for GSUS inflammation may be too sensitive when evaluated in RA patients and non-arthritic controls (44-46). PD signal may also be present in healthy subjects (47) and the PD sensitivity for low flows in small vessels, which are most relevant to rheumatological US assessing (subclinical) inflammation, may be highly variable for different US machines (48, 49). These indicate that there is still room for improvements in the domain of the validity of US.

Also regarding reliability, issues remain. OMERACT will probably publish a general scoring system (50) which would increase standardized scanning since over the years, several groups have used various US scoring systems, including several definitions and cut offs for inflammation, even differing in quantitative and semiquantitative measuring (42, 51-54). Also the scanning methodology may need more standardization, despite the publication of European league against rheumatism (EULAR) recommendations on how to scan (55). So there is still room for improvement, especially in the domain of acquisition reliability, i.e. dynamic scanning and scoring of patients by more than one observer (56, 57). A general scoring system by the OMERACT would increase standardisation, but would also necessitate new reliability exercises.

Apart from all this, reliability statistics themselves could, and should be, topic of debate. The reliability statistics used most frequently, i.e. kappa and intraclass correlation coefficient, aim to correct for agreement by chance alone. They are, however, largely dependent on the number and distribution of abnormalities scored, in contrast to the actual agreement between observers (58-60).

Despite the remaining issues, thanks to OMERACT, US became a feasible and accepted tool in rheumatology. We have learned that US has the potential to predict progression to RA from undifferentiated arthritis (UA) (12, 13), which would potentially increase detection of new RA patients early, making it possible to initiate treatment even earlier in the 'window of opportunity', possibly altering outcomes of patients beneficially. Inflammation detected at an US examination is also associated with erosive progression at joint level in response to treatment studies (61), which indicates that US could be a tool to discriminate patients who need more intensive treatment than others. US is also able to predict flares and erosive progression in RA patients in remission (14-17), which would make it a tool to discriminate patients who can taper their medication from patients who cannot or even need more intensive treatment. For these reasons US has become part of clinical decision making, with the suggestion to even include US in remission criteria (62, 63) and classification criteria for RA (9).

However, the development of US in RA has been a simultaneous development with the 'treat to target' and 'tight control' paradigm shifts in RA. This means that, apart from the issues still remaining as described above, it may be necessary to re-evaluate the place of US in RA patients treated according these current paradigms.

OUTLINE OF THIS THESIS

The development of US in rheumatology is an ongoing process, however there is still room for improvement (26). Despite the progress that has been made this mil-

1

lennium, technical considerations still exist in rheumatological US which are briefly described in the introduction of this thesis. One of these considerations lies in the domain of validity of both GSUS and PDUS. Therefore, in **chapter 2** a study is described in which a US anechogenicity is investigated as a possible explanation of the false positives in GSUS scanning. Tenosynovitis has been studied less than joint synovitis but its assessment may have to deal with similar validity issues. In **chapter 3** the value of current US definitions for tenosynovitis is investigated in a cohort of clinically idiopathic CTS patients. The ultrasonographic findings are validated with surgical, peroperative evaluation and histological examination. We evaluated the validity issues in the domain of PDUS by building a flow phantom on which the performance of the PD modality of several US machines is compared; the results are presented in **chapter 4**.

Interobserver reliability remains an issue in rheumatological US (56, 57). This may be because of the lack of clear definitions, nature of the technique and scanned structures, but also due to the inherent characteristics of reliability statistics itself. **Chapter 5** begins with a short introduction regarding this issue, followed by a study that aimed to increase the reliability of the US examination of the wrist, a notoriously difficult joint to assess at US.

Part 2, clinical considerations in rheumatological US, begins with a short chapter on the possible issues with the validity of current clinical measures (**chapter 6**). The thesis proceeds with a systematic review (**chapter 7**) to investigate the current status of the role of US in diagnosing RA and evaluating remission in RA. Important clinical issues remain, i.e. investigating the use of midrange equipment in multicenter studies (64), especially, in the light of the new treatment paradigms associated with better treatment options of RA, including biological treatment. It was evaluated what the value of US added to standard clinical assessments is for the prediction of clinical and radiological outcomes in RA patients (64). This is described in **chapter 8**, describing follow-up of a cohort of newly diagnosed RA patients for 1 year to investigate the added, predictive value of US for failing to reach DAS28 remission. **Part 2** closes with a study (**chapter 9**) in which US is used as a reference standard to develop an algorithm for a new imaging modality. This thesis is concluded with a general discussion including recommendations for a future research agenda.

The aims of this thesis can be summarized as:
1. To increase the validity of GSUS and PDUS
2. To increase the reliability of GSUS and PDUS for wrists
3. To evaluate the added, predictive value of US in patients with active RA

Note: Abbreviations are explained in the abbreviations list in the appendix.

REFERENCES

1. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. Arthritis Res. 2002;4 Suppl 3:S265-72.
2. http://www.nationaalkompas.nl/gezondheid-en-ziekte/ziekten-en aandoeningen/bewegingsstel-sel-en-bindweefsel/reumatoide-artritis-ra/omvang/.
3. Rothschild BM, Turner KR, DeLuca MA. Symmetrical erosive peripheral polyarthritis in the Late Archaic Period of Alabama. Science. 1988;241:1498-501.
4. Aceves-Avila FJ, Medina F, Fraga A. The antiquity of rheumatoid arthritis: a reappraisal. J Rheumatol. 2001;28:751-7.
5. Storey GD. Alfred Baring Garrod (1819-1907). Rheumatology (Oxford). 2001;40:1189-90.
6. Schlosstein L, Terasaki PI, Bluestone R, Pearson CM. High association of an HL-A antigen, W27, with ankylosing spondylitis. N Engl J Med. 1973;288:704-6.
7. Brewerton DA, Hart FD, Nicholls A, Caffrey M, James DC, Sturrock RD. Ankylosing spondylitis and HL-A 27. Lancet. 1973;1:904-7.
8. van der Helm-van Mil AH, Huizinga TW. Advances in the genetics of rheumatoid arthritis point to subclassification into distinct disease subsets. Arthritis Res Ther. 2008;10:205.
9. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis. 2010;69:1580-8.
10. Wakefield RJ, D'Agostino MA, Naredo E, Buch MH, Iagnocco A, Terslev L, et al. After treat-to-target: can a targeted ultrasound initiative improve RA outcomes? Postgrad Med J. 2012;88:482-6.
11. Jacobs JW, Ten Cate DF, van Laar JM. Monitoring of rheumatoid arthritis disease activity in individual patients: still a hurdle when implementing the treat-to-target principle in daily clinical practice. Rheumatology (Oxford). 2014.
12. 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.
13. 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.
14. 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.
15. 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.
16. Peluso G, Michelutti A, Bosello S, Gremese E, Tolusso B, Ferraccioli G. Clinical and ultrasonographic remission determines different chances of relapse in early and long standing rheumatoid arthritis. Ann Rheum Dis. 2011;70:172-5.
17. 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.

18. Manbachi A, Cobbold RSC. Development and application of piezoelectric materials for ultrasound generation and detection. Ultrasound. 2011;19:187-96.

19. Mason W.P. and Thurston R.N. eds.: *Physical Acoustics* (XV) AP, New York, 1981.43.

20. Donald I. Apologia: how and why medical sonar developed. Ann R Coll Surg Engl. 1974;54:132-40.

21. Fry WJ, Meyers R. Ultrasonic method of modifying brain st ructures. Confin Neurol. 1962;22:315-27.

22. Hauser I. DW, Frey R. and Ueberle A. . Experimentelle Untersuchungen fiber die Ultraschallwirkung auf das Jensen-Sarkom der Ratte. Zeitschrift ftir Krebsforsehung. 1949;56:449--81.

23. Mason W.P. and Thurston R.N. eds.: *Physical Acoustics* (XV) AP, New York, 1981.79.

24. Wild JJ, Reid JM. Application of echo-ranging techniques to the determination of structure of biological tissues. Science. 1952;115:226-30.

25. Donald I, Macvicar J, Brown TG. Investigation of abdominal masses by pulsed ultrasound. Lancet. 1958;1:1188-95.

26. Grassi W, Filippucci E. A brief history of ultrasound in rheumatology: where we were. Clin Exp Rheumatol. 2014;32:S3-6.

27. Szabo T.L. *Diagnostic Ultrasound Imaging -- Inside out,* Academic Press, Boston, 2014; 468.

28. Rubin JM, Adler RS, Fowlkes JB, Spratt S, Pallister JE, Chen JF, et al. Fractional moving blood volume: estimation with power Doppler US. Radiology. 1995;197:183-90.

29. Torp-Pedersen ST, Terslev L. Settings and artefacts relevant in colour/power Doppler ultrasound in rheumatology. Ann Rheum Dis. 2008;67:143-9.

30. Bude RO, Rubin JM. Power Doppler sonography. Radiology. 1996;200:21-3.

31. Yoon DY, Choi BI, Kim TK, Han JK, Yeon KM. Influence of instrument settings on flow signal and background noise in power Doppler US. An experimental study using a flow phantom with hyperechoic background. Invest Radiol. 1999;34:781-4.

32. Elfarnawany M, Pinter SZ, Lacefield JC. Improved objective selection of power Doppler wall-filter cut-off velocity for accurate vascular quantification. Ultrasound Med Biol. 2012;38:1429-39.

33. McDonald DG, Leopold GR. Ultrasound B-scanning in the differentiation of Baker's cyst and thrombophlebitis. Br J Radiol. 1972;45:729-32.

34. Cooperberg PL, Tsang I, Truelove L, Knickerbocker WJ. Gray scale ultrasound in the evaluation of rheumatoid arthritis of the knee. Radiology. 1978;126:759-63.

35. Grassi W, Tittarelli E, Pirani O, Avaltroni D, Cervini C. Ultrasound examination of metacarpophalangeal joints in rheumatoid arthritis. Scand J Rheumatol. 1993;22:243-7.

36. Newman JS, Adler RS, Bude RO, Rubin JM. Detection of soft-tissue hyperemia: value of power Doppler sonography. AJR Am J Roentgenol. 1994;163:385-9.

37. Wakefield RJ, Green MJ, Marzo-Ortega H, Conaghan PG, Gibbon WW, McGonagle D, et al. Should oligoarthritis be reclassified? Ultrasound reveals a high prevalence of subclinical disease. Ann Rheum Dis. 2004;63:382-5.

38. Wakefield RJ, Gibbon WW, Conaghan PG, O'Connor P, McGonagle D, Pease C, et al. The value of sonography in the detection of bone erosions in patients with rheumatoid arthritis: a comparison with conventional radiography. Arthritis Rheum. 2000;43:2762-70.

39. Szkudlarek M, Court-Payen M, Strandberg C, Klarlund M, Klausen T, Ostergaard M. Power Doppler ultrasonography for assessment of synovitis in the metacarpophalangeal joints of patients with rheumatoid arthritis: a comparison with dynamic magnetic resonance imaging. Arthritis Rheum. 2001;44:2018-23.

40. Hau M, Kneitz C, Tony HP, Keberle M, Jahns R, Jenett M. High resolution ultrasound detects a decrease in pannus vascularisation of small finger joints in patients with rheumatoid arthritis

receiving treatment with soluble tumour necrosis factor alpha receptor (etanercept). Ann Rheum Dis. 2002;61:55-8.

41. http://www.omeract.org/index.html.

42. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostino MA, et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol. 2005;32:2485-7.

43. Wakefield RJ, D'Agostino MA, Iagnocco A, Filippucci E, Backhaus M, Scheel AK, et al. The OMERACT Ultrasound Group: status of current activities and research directions. J Rheumatol. 2007;34: 848-51.

44. Witt M, Mueller F, Nigg A, Reindl C, Leipe J, Proft F, et al. Relevance of grade 1 gray-scale ultrasound findings in wrists and small joints to the assessment of subclinical synovitis in rheumatoid arthritis. Arthritis Rheum. 2013;65:1694-701.

45. Balsa A, de Miguel E, Castillo C, Peiteado D, Martin-Mola E. Superiority of SDAI over DAS-28 in assessment of remission in rheumatoid arthritis patients using power Doppler ultrasonography as a gold standard. Rheumatology (Oxford). 2010;49:683-90.

46. van de Stadt LA, Bos WH, Meursinge Reynders M, Wieringa H, Turkstra F, van der Laken CJ, et al. The value of ultrasonography in predicting arthritis in auto-antibody positive arthralgia patients: a prospective cohort study. Arthritis Res Ther. 2010;12:R98.

47. Kitchen J, Kane D. Greyscale and power Doppler ultrasonographic evaluation of normal synovial joints: correlation with proand anti-inflammatory cytokines and angiogenic factors. Rheumatology (Oxford). 2014.

48. Browne JE, Watson AJ, Hoskins PR, Elliott AT. Validation of a sensitivity performance index test protocol and evaluation of colour Doppler sensitivity for a range of ultrasound scanners. Ultrasound Med Biol. 2004;30:1475-83.

49. Koski JM, Saarakkala S, Helle M, Hakulinen U, Heikkinen JO, Hermunen H. Power Doppler ultrasonography and synovitis: correlating ultrasound imaging with histopathological findings and evaluating the performance of ultrasound equipments. Ann Rheum Dis. 2006;65:1590-5.

50. D'Agostino MA, Conaghan PG, Naredo E, Aegerter P, Iagnocco A, Freeston JE, et al. The OMERACT ultrasound task force -Advances and priorities. J Rheumatol. 2009;36:1829-32.

51. Backhaus M, Ohrndorf S, Kellner H, Strunk J, Backhaus TM, Hartung W, et al. Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. Arthritis Rheum. 2009;61: 1194-201.

52. Naredo E, Bonilla G, Gamero F, Uson J, Carmona L, Laffon A. Assessment of inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with grey scale and power Doppler ultrasonography. Ann Rheum Dis. 2005;64:375-81.

53. Scheel AK, Hermann KG, Kahler E, Pasewaldt D, Fritz J, Hamm B, et al. A novel ultrasonographic synovitis scoring system suitable for analyzing finger joint inflammation in rheumatoid arthritis. Arthritis Rheum. 2005;52:733-43.

54. Szkudlarek M, Court-Payen M, 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.

55. Backhaus M, Burmester GR, Gerber T, Grassi W, Machold KP, Swen WA, et al. Guidelines for musculoskeletal ultrasound in rheumatology. Ann Rheum Dis. 2001;60:641-9.

56. Cheung PP, Dougados M, Gossec L. Reliability of ultrasonography to detect synovitis in rheumatoid arthritis: a systematic literature review of 35 studies (1,415 patients). Arthritis Care Res (Hoboken). 2010;62:323-34.

57. Cheung PP, Kong KO, Chew LC, Chia FL, Law WG, Lian TY, et al. Achieving consensus in ultrasonography synovitis scoring in rheumatoid arthritis. Int J Rheum Dis. 2013.

58. Ten Cate DF, Luime JJ, Hazes JM, Jacobs JW, Landewe R. Does the intraclass correlation coefficient always reliably express reliability? Comment on the article by Cheung et al. Arthritis Care Res (Hoboken). 2010;62:1357-8; author reply 8.

59. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol. 1990;43:543-9.

60. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. J Clin Epidemiol. 1990;43:551-8.

61. Fukae J, Isobe M, Kitano A, Henmi M, Sakamoto F, Narita A, et al. Positive synovial vascularity in patients with low disease activity indicates smouldering inflammation leading to joint damage in rheumatoid arthritis: time-integrated joint inflammation estimated by synovial vascularity in each finger joint. Rheumatology (Oxford). 2013;52:523-8.

62. Haavardsholm EA, Lie E, Lillegraven S. Should modern imaging be part of remission criteria in rheumatoid arthritis? Best Pract Res Clin Rheumatol. 2012;26:767-85.

63. Ostergaard M, Moller-Bisgaard S. Rheumatoid arthritis: Is imaging needed to define remission in rheumatoid arthritis? Nat Rev Rheumatol. 2014;10:326-8.

64. Mandl P, Kurucz R, Niedermayer D, Balint PV, Smolen JS. Contributions of ultrasound beyond clinical data in assessing inflammatory disease activity in rheumatoid arthritis: current insights and future prospects. Rheumatology (Oxford). 2014.

PART 1

Fundamental domain
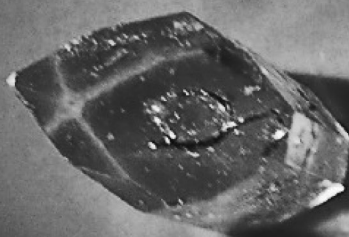
CHAPTER 2
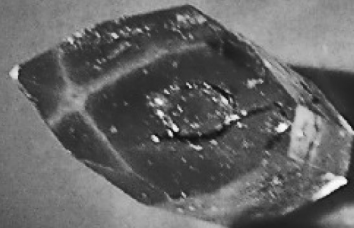
Is the frequent sonographic anechoic area distally in metacarpophalangeal joints a sign of arthritis?

Ten Cate DF, Luime JJ, Hazes JM, Kleinrensink GJ, Jacobs JW.

ABSTRACT

In clinical practice, ultrasonography (US) often reveals, in the dorsal scan, a small anechoic area distally in both inflamed and clinically non-inflamed metacarpophalangeal joints. This ''distal anechogenicity in the metacarpophalangeal joint'' (DAEM) might thus be scored false positively as arthritis. We aimed to investigate whether the DAEM is a sign of arthritis. We evaluated the prevalence of DAEMs in 24 non-arthritic subjects. We then compared the dimensions of the DAEM in 10 non-arthritic subjects with a DAEM and 7 consecutive rheumatoid arthritis (RA) outpatients, using 2-D and 3-D ultrasound. Furthermore, we dissected two fresh-frozen postmortem hand specimens after US. A DAEM was observed in the metacarpophalangeal 2 (MCP2) joints of 54% of the 24 non-selected non-arthritic individuals; in none of those did the joint exhibit a power Doppler signal. A DAEM was observed in 86% of the 7 RA patients. Dimensions of DAEMs did not statistically significantly differ between these groups. At 3-D imaging and dissection, the DAEM was found to be an extension of the metacarpophalangeal joint capsule. In conclusion, DAEMs occur frequently and are not a sign of arthritis, but are distal joint recesses. This should be taken into account when using current sensitive ultrasonographic scoring systems grading arthritis.

INTRODUCTION

Ultrasonography (US) is used increasingly in rheumatology for diagnostic purposes and to evaluate disease activity, response to treatment and prognosis (1-4). Inclusion of US findings in remission criteria (5) and use of US findings to confirm clinical findings using the rheumatoid arthritis (RA) classification criteria (6) have been suggested. At US, clinically relevant subclinical disease activity may be detected (7,8) in arthritis patients in clinical remission (9) and in clinically non-inflamed joints (3). Grading of disease activity to make important decisions regarding therapy, especially in treat-to-target strategies in RA, is dependent on correct assessment of synovitis. US can be used as a tool for this purpose only if it validly can confirm and grade or exclude synovitis.

According to current insights, anechogenicity and hypo-echogenicity within joints at gray-scale US (GSUS) represent synovial fluid (hydrops) and synovial hypertrophy, respectively, which may be signs of arthritis, especially in combination with positive power Doppler (PD) signals (10). However, scanning joints in non-arthritic individuals using GSUS scoring for synovitis may result in a large number of false positive mild synovitis scores (11,12).

In daily clinical practice, when scanning the dorsal aspect of the joint, often just distally in the joint surface of non-inflamed metacarpophalangeal (MCP) joints, a small hypo-echoic to anechoic area is seen at the base of the proximal phalanx (Fig. 1). This ''distal anechogenicity in the metacarpophalangeal joint'' (DAEM) is defined as a compressible hypo-echoic to anechoic area in the joint space from the proximal bor-
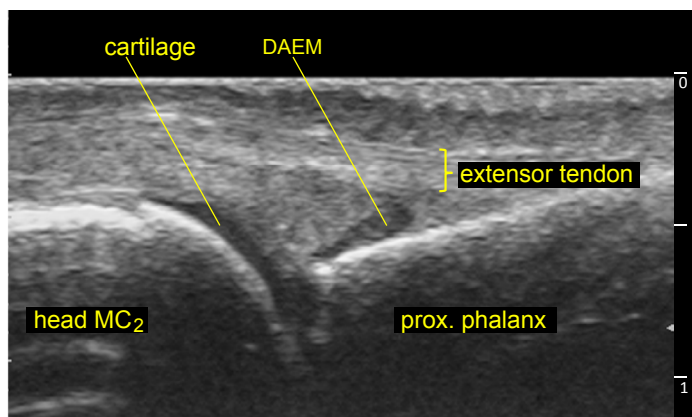
Figure 1. Longitudinal dorsal ultrasonographic scan of a second metacarpophalangeal (MCP2) joint of a non-arthritic subject with a ''distal anechogenicity in the metacarpophalangeal joint'' (DAEM). head MC2 = head of second metacarpal bone; prox. phalanx = proximal phalanx; DAEM = distal anechogenicity in the metacarpophalangeal joint.

der of the proximal phalanx distally toward where the anechogenicity ceases (see Figs. 1 and 2). In our experience it occurs predominantly in MCP2 joints, but can be found in all MCP joints. Frequent occurrence in non-inflamed joints suggests a physiologic rather than a pathologic entity. However, if the commonly used criterion of hypo- or anechogenicity indicative of joint inflammation is applied, a DAEM could be scored as inflamed (10). This could result in incorrect assessments of disease activity, for instance, falsely reclassifying an RA patient in full remission as a patient with RA not in full remission. This could lead to unnecessary therapeutic actions when treating to
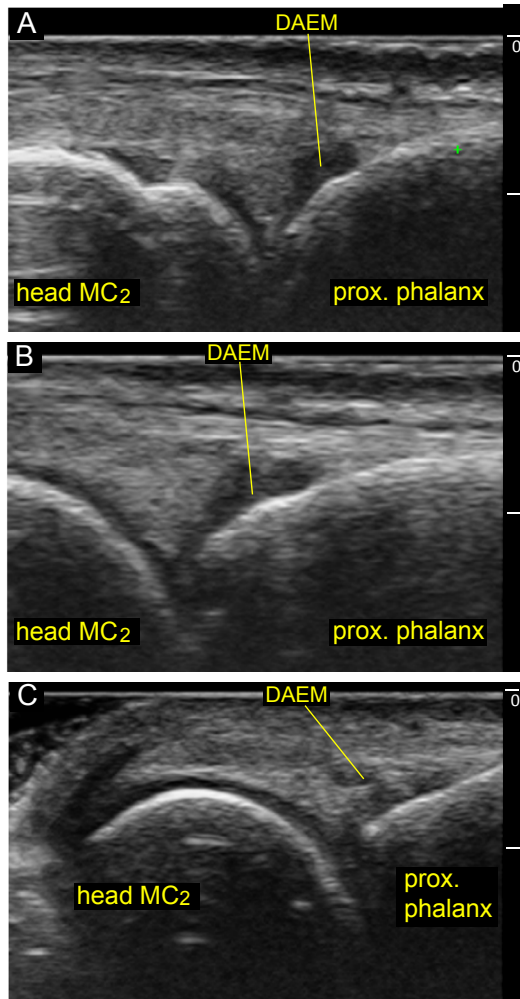
Figure 2. Dynamic longitudinal 2-D ultrasonographic scan of a distal anechogenicity in the metacarpophalangeal joint (DAEM). (a) Neutral extended position of the joint. (b) Joint in 45° flexion. (c) Joint in full flexion. head MC2 = head of second metacarpal bone; prox. phalanx = proximal phalanx; DAEM = distal anechogenicity in the metacarpophalangeal joint.

target, if the target is absence of synovitis. For these reasons, we deemed it important to investigate DAEMs in depth, which has not been done before.

The aim of this study was to roughly estimate the prevalence of DAEMs in non-arthritic subjects and RA patients and to establish whether the DAEM is an ultrasonographic artifact, normal physiology or an inflammatory entity, that is, sign of arthritis.

2

METHODS

In the first part of the study, both MCP2 joints of 24 non-arthritic consecutive volunteers without a history of joint signs or symptoms were scanned to evaluate the prevalence of DAEMs. In the second part of the study, we sought to identify whether the DAEM is an ultrasonographic artifact, normal physiology or an inflammatory entity, that is, sign of arthritis. Therefore, we scanned both MCP2 joints of 10 subjects without a history of joint signs or symptoms, but with a known DAEM, and 7 consecutive RA patients. Subjects with a DAEM but no history of symptoms were identified from a convenient pool of students participating in didactic training sessions in our center (Erasmus University Medical Center). The RA patients were recruited from the rheumatology outpatient clinic of our center.

We used 2-D US to assess dimensions of DAEMs, dynamic 2-D US to assess how the dimensions of the DAEM change when the joint moves and 3-D US to assess anatomy. Also, two fresh-frozen postmortem hand specimens were scanned and dissected after US and injection of the MCP2 joint with Fillopaq (Anubifix, The Netherlands) to further investigate the anatomy of DAEMs.

The study protocol was approved by the local ethics committee. All subjects gave their written informed consent before the US examination.

Ultrasonography

Both MCP2 joints of all study subjects were scanned by an experienced ultrasonographer (D.T.C.), certified by the Dutch College of Rheumatology. Only MCP2 joints were evaluated, although in our experience, DAEMs may be found in other MCP joints as well. The US investigations comprised both GSUS and power Doppler US (PDUS). For GSUS, we used an Aloka α7 (Aloka, Tokyo, Japan) with a 5 to 16-MHz compound array UST-5411 transducer. For the 3-D images, we used the GE Logiq E9 (GE Healthcare, Buckinghamshire, UK) machine with an RSP 6-16 transducer at 15 MHz. For PD examination, the color gain was set just at the disappearance of color noise; the

velocity range was 1.2 cm/s for the Aloka, and the pulse repetition frequency was 1.1 kHz for the GE. The frequency was set at 7.27 MHz for the Aloka and 8.3 MHz for the GE. Wall filters were set at low levels. The size and position of the color box were adjusted to include the subcutaneous tissue to recognize artifacts caused by blood vessels above the joint. All joints were scanned longitudinally from the dorsal side in neutral, extended position, according to European League Against Rheumatism guidelines (13).

In the first part of the study, we recorded an image of each MCP2 joint in neutral position. In the second part of the study, we investigated the functional dynamic properties of DAEMs, that is, how the dimensions of DAEMs change at different angles of the joint. To ascertain these angles, we scanned the joints additionally in 45° flexion, using a foam wedge to maintain this angle, and in maximal flexion. This resulted in 12 images per subject: six GSUS images and six 3-D US images. Each image included the head of the metacarpal bone, the proximal phalanx and the joint capsule. If a DAEM was present, its margins were traced, and the cross-sectional area and circumference were recorded.

For the 3-D analysis, the margins of the DAEM and entire joint space were traced in each of the 12 slices in the longitudinal plane within the scanned volume. In doing so, we also observed whether there was continuity between the DAEM and the joint space. Using VOCAL software, we calculated the volumes and created a 3-D image for both the DAEM and the joint space.

Postmortem hand specimens

Two fresh-frozen postmortem right-hand specimens with no macroscopically visible signs of (osteo)arthritis were scanned using the same scanning protocol as used in the second part of the study. To evaluate the shape of the distal joint space and its possible relationship to the DAEM, we injected the cadaver MCP2 joints with Fillopaq. This substance, when injected, has the viscosity of water at low temperatures and, when hardened at room temperature (takes about 10 min), the consistency of rubber. Thereafter, the joints were cut in half longitudinally using a Multimaster saw (FEIN, Schwaebisch Gmuend, Germany). The location of the DAEM was determined, and the presence of Fillopaq was assessed. We removed the Fillopaq and examined the shape of the joint. Photographs were taken of the procedure, and anatomic structures in the joint were identified (Fig. 3).
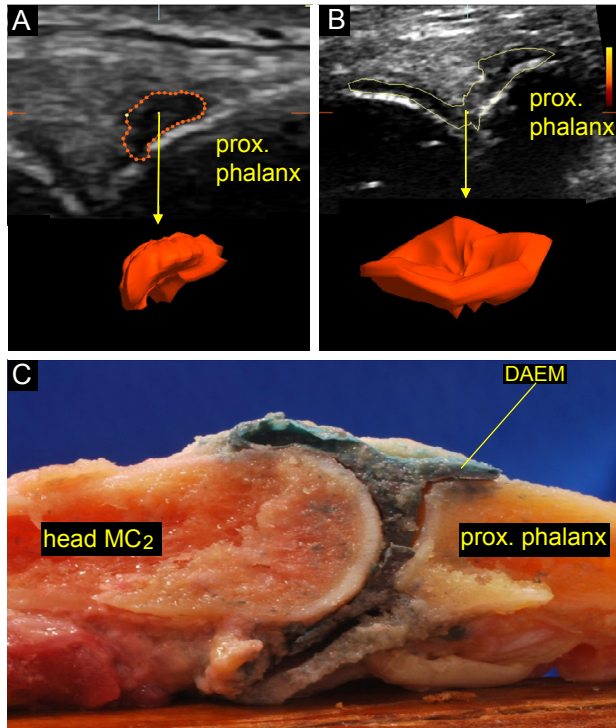
Figure 3. (a) Three-dimensional representation of a distal anechogenicity in the metacarpophalangeal joint. (b) Three-dimensional representation of the entire dorsal joint space. (c) Photograph of a postmortem hand specimen with the joint space filled with Fillopaq, illustrating continuity with the distal anechogenicity in the metacarpophalangeal joint. head MC2 = head of second metacarpal bone; prox. phalanx = proximal phalanx; DAEM = distal anechogenicity in the metacarpophalangeal joint.

Statistics

Differences between group medians for DAEM dimensions were tested for statistical significance with the Mann–Whitney *U*-test or Wilcoxon signed-rank test, where appropriate. Differences between proportions were tested for significance using Fisher's exact test, where *p*< 0.05 was considered to indicate significance. For analyses, Number Cruncher Statistical Software 2007 (Kaysville, UT, USA) was used.

RESULTS

In 13 of the 24 (54%) consecutively recruited nonarthritic controls (median age: 30.5 y, range: 22–64 y, 9 females and 4 males), a DAEM was present in at least one MCP2

joint. There was a DAEM in at least one MCP2 joint in 6 of the 7 (86%) non-selected consecutive RA outpatients (median age: 55, range: 52–71 y, 6 females and 1 male, median disease duration: 4 y, range: 1.5–13 y). The prevalence of DAEM did not significantly differ between consecutively recruited controls and RA patients ($p >0.2$). The RA patients generally were older than the controls. Among the 7 oldest controls (median age: 50 y, range: 40–64 y), 3 had a DAEM (43%), and among the other 17 controls (median age: 28 y, range: 22–38 y), 10 (59%) had a DAEM, indicating that the prevalence of DAEMs does not seem to be higher in the older age group. A DAEM was also present in the MCP2 joint of both postmortem specimens. None of the asymptomatic individuals developed clinical arthritis during a follow-up period of 15 months.

No significant differences were found in cross-sectional area and circumference of the DAEM between control subjects and RA patients with a DAEM; median (range) cross-sectional areas were 2 (1–5) and 2 (1–10) mm2, respectively ($p = 0.38$), and median (range) circumferences were 9.0 (0.52–1.36) and 9.5 (0.7–1.54) mm, respectively ($p = 0.98$).

Analysis of DAEMs in various degrees of flexion of the MCP joint was performed on the right hand of controls and patients (total n = 16). The maximum flexed position had a median of 88° (range: 55°–95°) across all subjects. When the joint was moved from full extension to maximum flexion, DAEMs remained visible in all subjects, but median cross-sectional area significantly decreased from 2.0 to 1.5 mm2 ($p <0.05$) (Fig. 2).

Clear synovitis, defined as elevation of the joint capsule above the joint and a PD signal within the joint, was observed in three right-sided and two left-sided MCP2 joints of the 7 RA patients. No PD signal was found in any of the MCP2 joints of control subjects or in the MCP2 joints of 3 of the 7 RA patients. The median (range) cross-sectional area of the DAEM was 2 (range: 1–10) mm2 in the 3 right-sided MCP joints with synovitis versus 1.5 (range: 0–3) mm2 in the 4 right-sided MCP joints of RA patients without synovitis in this joint ($p = 0.59$).

Regarding the nature of the DAEM, we studied each slice of scanned volume separately. The DAEM was continuous with the joint space, indicating that the DAEM is a part of the joint space and not an ultrasonographic artifact. This was confirmed with the results at dissection of the MCP joints of the two specimens, as illustrated in Figure 3c. Figure 3 (a, b) comprises 3-D representations of the DAEM and of the entire dorsal joint space separately in one subject.

DISCUSSION

We observed a high prevalence of DAEMs in nonarthritic subjects and similar dimensions of DAEMs in non-arthritic subjects and RA patients. Furthermore, we noted an absence of PD signal within the joints of the non-arthritic subjects with a DAEM and an absence of arthritis in these individuals at follow-up. Together our findings suggest that DAEMs are not a sign of synovitis. Instead, the results of 3-D US and cadaver dissection suggest that DAEMs are distal joint recesses.

Distal recesses of MCP joints have been described before in brief (14,15). Our study is the first to investigate DAEMs in detail. DAEMs might be a cause of the high prevalence of mild synovitis scores in non-arthritic individuals (11,12), but this is merely a hypothesis.

The findings of our study could have consequences in the use of current sensitive scoring protocols. False interpretation of a DAEM appearing as a hypoechogenicity or anechogenicity of the joint as a sign of (mild/early) synovitis may have clinical implications, especially in treat-to-target strategies when the target is remission. Furthermore, it could jeopardize the added value of US in current and future RA classification criteria and remission criteria for RA.

The primary limitation of this study is its modest sample size. Despite this, the high frequency of DAEMs in both control subjects and patients, in addition to the results of the 3-D and cadaver analyses, supports our hypothesis that the presence of a DAEM itself is not directly related to arthritis. Control subjects and arthritis patients differed in age, but the prevalence of DAEMs did not seem to vary when the control group was subdivided into age groups. The trends in DAEMs between patients and controls of various ages indicated by our data should be further explored and validated in larger samples.

One could comment that injection of Fillopaq artificially distended the joint cavity. However, the shape of the joint capsule was still rather concave after injection. Moreover, the dissection results were in line with the findings at 3-D scanning. We did not assess the occurrence of DAEMs in joints other than MCP2. However, our study clearly indicates that DAEMs are not a rare phenomenon. This emphasizes that anatomic studies and ultrasound studies in non-arthritic subjects are needed to validate US findings, to increase the value of US in clinical rheumatology.

CONCLUSIONS

A DAEM is not an ultrasonographic artifact or a sign of arthritis, but is a distal joint recess commonly found in non-arthritic subjects and RA patients. The DAEM may jeopardize sensitive US scoring systems grading synovitis.

ACKNOWLEDGMENTS
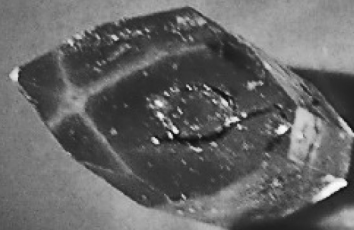
Note: Abbreviations are explained in the abbreviations list in the appendix.

REFERENCES

1. Brown AK, Quinn MA, Karim Z, Conaghan PG, Peterfy CG, Hensor E et al. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying anti-rheumatic drug-induced clinical remission: Evidence from an imaging study may explain structural progression. Arthritis Rheum 2006;54:3761–3773.

2. 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;48:1092–1097.

3. Wakefield RJ, Green MJ, Marzo-Ortega H, Conaghan PG, Gibbon WW, McGonagle D et al. Should oligoarthritis be reclassified? Ultrasound reveals a high prevalence of subclinical disease. Ann Rheum Dis 2004;63:382–385.

4. Ziswiler HR, Aeberli D, Villiger PM, Moller B. High-resolution ultrasound confirms reduced synovial hyperplasia following rituximab treatment in rheumatoid arthritis. Rheumatology (Oxford) 2009; 48:939–943.

5. Haavardsholm EA, Lie E, Lillegraven S. Should modern imaging be part of remission criteria in rheumatoid arthritis? Best Pract Res Clin Rheumatol 2012;26:767–785.

6. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd, et al. 2010 rheumatoid arthritis classification criteria: An American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2010;69:1580–1588.

7. 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;58:2958–2967.

8. 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.

9. Saleem B, Brown AK, Keen H, Nizam S, Freeston J, Wakefield R, Karim Z, Quinn M et al. Extended report: Should imaging be a component of rheumatoid arthritis remission criteria? A comparison between traditional and modified composite remission scores and imaging assessments. Ann Rheum Dis 2011;70:792–798.

10. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostino MA et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol 2005;32:2485–2487.

11. Ellegaard K, Torp-Pedersen S, Holm CC, Danneskiold-Samsoe B, Bliddal H. Ultrasound in finger joints: Findings in normal patients and pitfalls in the diagnosis of synovial disease. Ultraschall Med 2007;28:401–408.

12. Szkudlarek M, Wakefield RJ, Backhaus M, Terslev L. The discriminatory capacity of ultrasound in rheumatoid arthritis: Active vs inactive, early vs advanced, and more. Rheumatology (Oxford) 2012; 51(Suppl 7):vii6–vii9.

13. Backhaus M, Burmester GR, Gerber T, Grassi W, Machold KP, Swen WA et al. Working Group for Musculoskeletal Ultrasound in the ESCoICSiTT. Guidelines for musculoskeletal ultrasound in rheumatology. Ann Rheum Dis 2001;60: 641–649.

14. Bade H, Koebke J, Nieden A. Radiologic anatomy of the metacarpophalangeal joints II to V. Surg Radiol Anat 1997;19:323–327.

15. McNally EG. Ultrasound of the small joints of the hands and feet: Current status. Skeletal Radiol 2008;37:99–113.

2

CHAPTER 3

A comparison between ultrasonographic, surgical and histological assessment of tenosynovitis in a cohort of idiopathic carpal tunnel syndrome patients.

Ten Cate DF, Glaser N, Luime JJ, Lam KH, Jacobs JW, Selles R, Hazes JM, Bertleff M

ABSTRACT

Carpal tunnel syndrome (CTS) may be caused by subclinical tenosynovitis which may be detected by ultrasonography (US). The objective of this study is to investigate whether ultrasonography has a place in the workup of idiopathic CTS patients. Therefore, we investigated the prevalence of tenosynovitis and its association with the clinical outcome of surgery. A cohort of 31 consecutive idiopathic CTS patients (33 wrists) who were a candidate for carpal tunnel release (CTR) surgery was assessed using greyscale ultrasonography (GSUS) and power Doppler ultrasonography (PDUS). Peroperatively, tenosynovitis was evaluated macroscopically by the surgeon. Tissue samples from areas macroscopically suspected for tenosynovitis were taken for histological evaluation. The clinical outcome of the operation was assessed after 6 months and if applicable alternative diagnoses for the CTS were proposed. US tenosynovitis (OMERACT) was detected preoperatively in 58% of the wrists. Peroperatively, macroscopic tenosynovitis was detected visually in 88% of the wrists. Histological evaluation demonstrated a limited influx of lymphocytes indicative of a mild chronic inflammatory response in 19%. Non-specific reactive changes were observed in 78% of the cases. Ultrasonographically defined tenosynovitis was associated with an OR of 2.81 (95% CI 0.61–13) for responding well to surgery. Most cases of ultrasonographic and peroperatively defined tenosynovitis were classified by histology as reactive changes. The presence of ultrasonographic tenosynovitis might be associated with a better clinical outcome of surgery.

INTRODUCTION

Carpal tunnel syndrome (CTS) is a common neuropathy of the median nerve, with reported prevalences ranging from 2.7 to 16% (1–3). In many cases, no underlying condition can be diagnosed, rendering it idiopathic. However, CTS is associated with rheumatoid arthritis (RA) and other inflammatory arthropathies, trauma, diabetes, acromegaly, hypothyroidism and pregnancy (4–7).

In this study, we investigated to what extent subclinical tenosynovitis is present in idiopathic CTS patients. Ultrasonography (US) is frequently used to detect and evaluate (subclinical) tenosynovitis, but there is little or no data on the correlation between surgical and histological findings regarding tenosynovitis in patients with tenosynovitis as diagnosed using US (8–10). Previous histological studies have reported that histological tenosynovitis is found in approximately 10 % of the idiopathic CTS patients (11, 12). In contrast, the prevailing opinion among hand surgeons is that macroscopic tenosynovitis is observed in over half the patients who undergo surgery for idiopathic CTS.

To investigate whether ultrasonography has a place in the workup of idiopathic CTS, the prevalence of tenosynovitis in idiopathic CTS patients was investigated using ultrasonography, surgery and histology. The association of ultrasonographic tenosynovitis with the clinical outcome of surgery was also evaluated.

MATERIAL AND METHODS

Patient characteristics

Between July and December 2011, we included 31 consecutive patients (33 wrists) with unilateral or bilateral idiopathic CTS who were nonresponsive to conservative night splint therapy and therefore were a candidate for open carpal tunnel release surgery. All patients were operated on in the department of hand surgery at the Xpert clinic in Rotterdam, the Netherlands. Exclusion criteria for our study were thyroid disease, ASA (American Society of Anaesthesiologists) classification 3 or more, and pregnancy. The study protocol was approved by the Medical Ethics Committee of Erasmus Medical Centre. Written informed consent was obtained from all patients prior to inclusion.

US evaluation

US examination was performed 1h prior to carpal tunnel release surgery. Patient position was according to EULAR guidelines (9), with the probe positioned proximal of the scaphoid and the pisiform bones, in longitudinal orientation relative to the median nerve.

All ultrasonography examinations were performed by a sonographer (NG), trained specifically for this occasion by an expert (DTC). The sonographer was blinded to the patients' clinical status. An Esaote MyLab5 with an LA435 probe (18–6 MHz) was used. Greyscale US (GSUS) settings were adjusted for each patient to obtain the optimal image. For Power Doppler US (PDUS), the frequency was set to the maximum of 8 MHz, with a pulse repetition frequency of 500 Hz. Low wall filters were used. The upper border of the box size was flush with the skin.

Tenosynovitis was defined by the OMERACT as hypoechoic or anechoic thickened tissue with or without fluid within the tendon sheath, seen in two perpendicular planes and possibly exhibiting a Doppler signal (see Fig. 1) (13). The extent of the hyperaemia was scored semi-quantitatively (0–3) as a percentage of vascularized tissue as detected on PDUS (14).

Peroperative evaluation

All patients were operated on by the same experienced hand surgeon (MB). During open carpal tunnel release surgery, the extent of the macroscopic tenosynovitis was recorded by the surgeon according to a three-point scale. As there is no generally accepted scoring system, this scale (see box 1) was devised prior to our study by five hand surgeons from the hand surgery clinic. A sample of tenosynovium was obtained when peroperative macroscopic tenosynovitis (grades 1 and 2) (see Fig. 2) was found. Samples were collected during surgery and stored in 4 % paraformaldehyde (PFA) in 0.1 M phosphate buffer (pH 7.1).

Box 1; Peroperative classification of flexor tenosynovitis of the wrist (See Fig. 2 also)

- Grade 0—normal tendon, no swelling
- Grade 1—minimal tenosynovitis (tenosynovial changes, with mild swelling of the tendons but no adhesions)
- Grade 2—extensive tenosynovitis (synovial changes, with swelling and adhesions of the tendons)
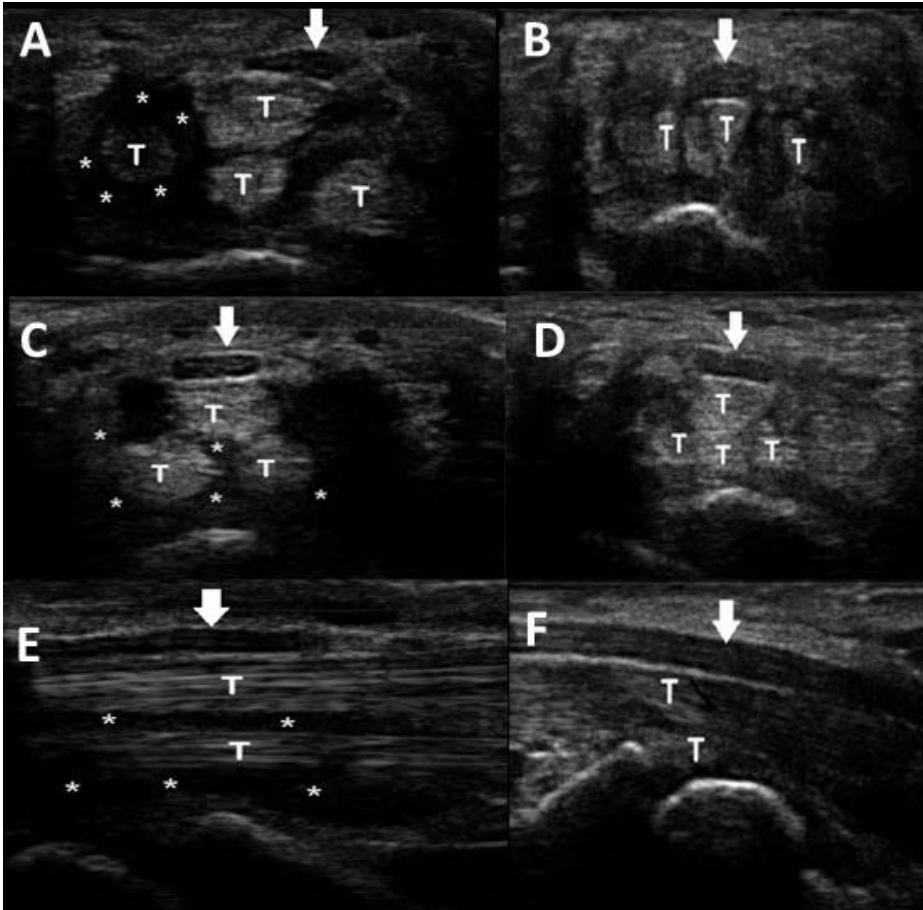
Figure 1 US images of the tenosynovium, tendons and median nerve at the carpal tunnel entrance. (a) US tenosynovitis of the flexor pollicis longus tendon in a 46-year-old male patient. Transverse US image of the tendons, tenosynovium and median nerve at the carpal tunnel inlet. Median nerve (*arrow*) is apparent as a speckled area immediately superficial to flexor tendons (T). Margins of the tendons and tenosynovium are marked (represented by an *asterisk*). (b) No US findings in a 48-year-old female patient. Transverse US image of the tendons, tenosynovium and median nerve at the carpal tunnel inlet. Median nerve (*arrow*) is apparent as a speckled area immediately superficial to flexor tendons (T). (c) US tenosynovitis of the flexor digitalis communis tendons in a 52-year-old female patient. Transverse US image of the tendons, tenosynovium and median nerve at the carpal tunnel inlet. Median nerve (*arrow*) is apparent as a speckled area immediately superficial to flexor tendons (T). Margins of the tendons and tenosynovium are marked (represented by an *asterisk*). (d) No US findings in a 57-year-old male patient. Transverse US image of the tendons, tenosynovium and median nerve at the carpal tunnel inlet. Median nerve (*arrow*) is apparent as a speckled area immediately superficial to flexor tendons (T). (e) US tenosynovitis in a 54-year-old female patient. Longitudinal US image of the tendons and tenosynovium with the median nerve as a point of orientation and the scaphoid as distal margin. Median nerve (*arrow*) is apparent as a hypoechoic structure immediately superficial to flexor tendons (T). Margins of the tendons and tenosynovium are marked (represented by an *asterisk*). (f) No US findings in a 44-year-old male patient. Longitudinal US image of the tendons and tenosynovium with the median nerve as a point of orientation and the scaphoid as distal margin. Median nerve (*arrow*) is apparent as a hypoechoic structure immediately superficial to flexor tendons (T)
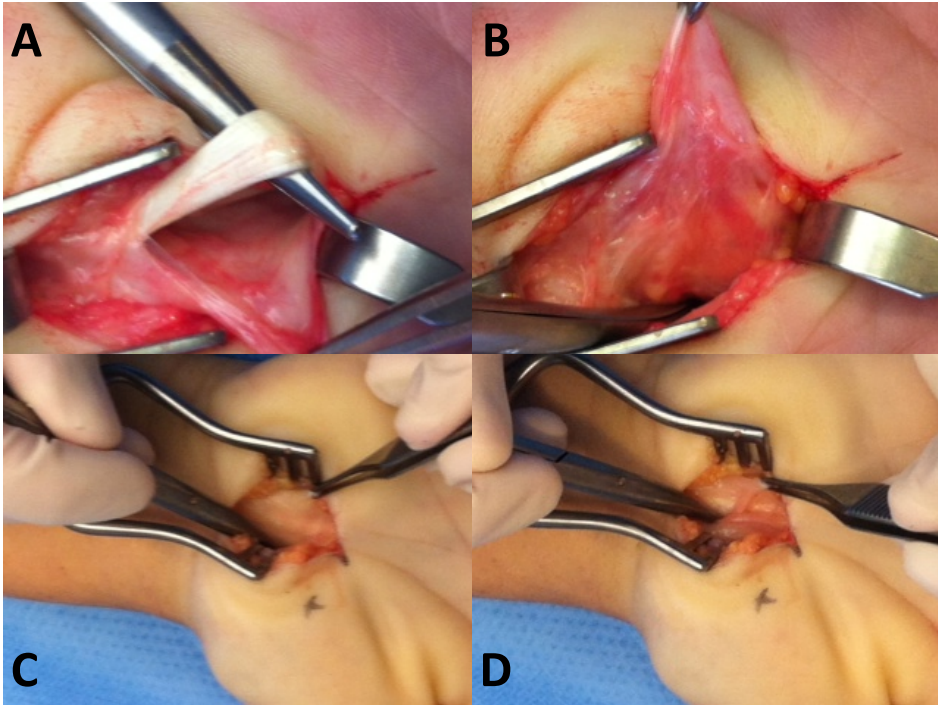
Figure 2 Surgical scoring system for tenosynovitis. (a) Grade 0—normal tendon, no swelling. (b) Grade 1— minimal tenosynovitis (tenosynovial changes, with mild swelling of the tendons without adhesions). (c + d) Grade 2— extensive tenosynovitis, swelling and adhesions of the tendons

----

Histological evaluation

The PFA-fixed material collected during surgery was dehydrated through a graded series of ethanol solutions in a critical point dryer. The material was then paraffin-embedded. Sections of 10-µm thickness were cut, mounted on glass slides and stained with haematoxylin and eosin. The pathologist (KL) evaluated the specimens and graded the histopathological changes in the tenosynovial tissue according to a protocol agreed upon prior to the study. (see box 2). The pathologist was also not aware of the surgical and ultrasonographic status of the patients. Based on expert opinion, grade 1 (no inflammation present) was judged not to be true tenosynovitis in

Box 2; Postoperative classification of flexor tenosynovitis of the wrist (See Fig. 3 for illustrations)

- Grade 0—no changes
- Grade 1—reactive changes, defined as an increased presence of fibroblasts
- Grade 2—tenosynovitis, defined as an influx of lymphocytes
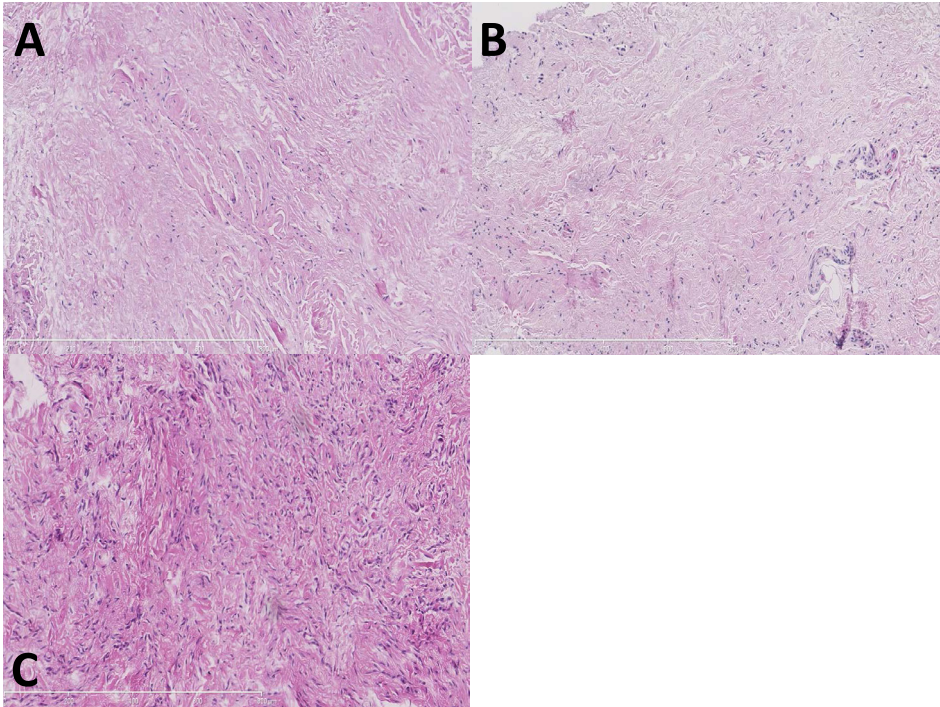
Figure 3 Histological scoring system for tenosynovitis. (a) Grade 0—no histopathological changes of tenosynovial tissue in a 61-year- old female patient. (b) Grade 1— reactive changes (increased fibroblast cellularity, but no inflammation) of tenosynovial tissue in a 68-year-old male patient. (c) Grade 2—minor tenosynovitis (mild inflammation, increased fibroblast cellularity and increased blood vessel proliferation) in tenosynovial tissue in a 74-year-old male patient

the analysis, whereas grade 2 (inflammation present) was judged to be true, though minor, tenosynovitis. (see Fig. 3).

Six-month postoperative evaluation

Patients were contacted 6 months postoperatively to assess the effect of the surgical intervention. The number of re-operations for CTS of the same wrist was recorded. A successful surgical intervention was defined as an absence of tingling sensations in the hand and fingers and if it was not necessary to wear a brace to relieve CTS symptoms 6 months after the operation. Any alternative diagnosis made by other specialists after surgery that could explain the CTS was also recorded (Table 1).

Statistical analysis

Data was analysed descriptively. Quantitative variables were presented as medians and ranges. Based on expert opinion, grades 1 and 2 in the surgical domain and grade 2 in the histological domain were considered to show presence of disease. The association of the presence of ultrasonographic tenosynovitis with the postoperative outcome was presented as a univariate odds ratio (OR) with a 95% confidence interval (CI). We could not calculate a required size of the study population since no previous ultrasound studies have been performed in this patient population to assess the amount of tenosynovitis. Statistical analysis was performed using Stata, version 12.0 (StataCorp, College station, TX, USA).

RESULTS

Patient characteristics

Thirty-one patients (33 wrists) were invited to participate in this study, and none refused. In two cases of surgical grade 1 tenosynovitis, biopsy specimens could not be obtained. Data is presented on the level of wrists. Table 1 shows patient characteristics. Nineteen of 33 wrists (58%) revealed tenosynovitis at GSUS (none showed PDUS), while 29 of 33 wrists (88%) revealed tenosynovitis peroperatively. Influx of lymphocytes, indicative of a chronic inflammatory response (grade 2), at histological examination was seen in 5 biopsies obtained in 27 wrists (19%). In addition, 21 of the biopsy specimens showed histologically reactive changes (grade 1). (Table 2) Four of the 19 operated wrists with tenosynovitis at GSUS were confirmed at histology.

Table 1. Characteristics of patients n=31 (median(range))

Age, years	54 (21-90)
Sex (n=31)	11 M* 20 F[#]
Duration of CTS symptoms in months	14 (4-72)
Wrists (%) with prior CTS surgery	12% one prior surgery; 6% two prior surgeries
Patients with an alternative diagnosis for CTS after 6 months (n=2)	Multiple Sclerosis (n=1) CTS-like symptoms due to cervical neuropathy (n=1)

*=Male, #=Female

Table 2. Prevalence of tenosynovitis findings per modality

Tenosynovitis Grade	US	Surgery	Histology
2	58	46	19
1	NA	42	78
0	42	12	3.7

For ultrasound, only a binary score was used: present (1) versus absent (0). For clarity, and consistency regarding the other modalities, the category "present" on ultrasound is defined as grade 2; NA=not applicable

3

Sixteen of these 19 wrists were also positive for surgical tenosynovitis. There was no agreement between modalities (see the Additional files for more information).

Six-month evaluation

At the 6-month evaluation, two patients had received an alternative diagnosis for the CTS symptoms. (Table 1) None of the 33 wrists required a reoperation within these 6 months. In 23 wrists (70 %), no tingling of the fingers was present or use of brace was necessary 6 months postoperatively. Four (21 %) of the 19 wrists in whom ultrasonographic tenosynovitis was detected did not improve, compared to 6 (43 %) of the 14 wrists without ultrasonographic tenosynovitis. From this follows that ultrasonographic tenosynovitis was associated with a favourable surgical outcome with a univariate OR of 2.81 (95 % CI 0.61–13).

DISCUSSION

To our knowledge, this is the first study to evaluate and compare US, surgical and histological findings in idiopathic CTS patients. Tenosynovitis was identified at GSUS and surgery in most of the operated wrists. Histology revealed in many of the biopsies that were taken reactive changes, but tenosynovitis—as defined by influx of lympho-cytes—was only detected in a few biopsies, similar to results of other studies (11, 12).

This would suggest that signs of tenosynovitis at ultrasound and surgery frequently are rather reactive changes, possibly as a response to mechanical stressors or pressure in the carpal tunnel, as has been hypothesized earlier (15). Although all patients were non-responsive to at least 6 weeks of conservative treatment, this may have been insufficient for several. These findings may question the use of glucocorticoid injections (16).

We found a trend suggesting that the presence of ultrasonographic tenosynovitis might be associated with a better clinical outcome of surgery. An explanation could be that surgical decompression of the carpal tunnel has an optimal symptomatic effect if indeed pressure is too high, e.g. by swelling of tissues or tenosynovitis. In the absence thereof, surgical decompression may be less useful.

This study has limitations. The US equipment was not high-end; the maximum frequency for the power Doppler modality was a modest 8 MHz. One could argue that this might account for the absence of PDUS signal. However, the absence of PDUS signal was consistent with the low prevalence of histological tenosynovitis in these patients. Another limitation is that two patients were included in the study with both wrists (paired data); however, an analysis without these wrists yielded similar results. Last, the patient number was modest, increasing the risk of a type-II statistical error, as was probably the case for the association between ultrasonographic tenosynovitis and a better clinical outcome of surgery. Nevertheless, based on this result, a prospective study incorporating ultrasonographical findings to develop an algorithm to treat idiopathic CTS patients seems justified.

CONCLUSIONS

In conclusion, at GSUS and at surgery, signs of tenosynovitis were detected in a large proportion of wrists. These tenosynovial changes were confirmed to be tenosynovitis by histology in a minority of the cases. The presence of ultrasonographic tenosynovial changes may be associated with a better clinical outcome of surgery. These findings could be the basis for a study investigating new treatment strategies for idiopathic CTS patients using an ultrasonographic examination.

Acknowledgments
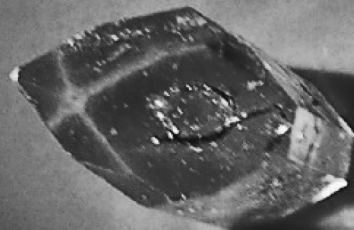
Note: Abbreviations are explained in the abbreviations list in the appendix.

REFERENCES

1. Atroshi I, Gummesson C, Johnsson R, Ornstein E, Ranstam J, Rosen I (1999) Prevalence of carpal tunnel syndrome in a general population. Jama 282:153–158
2. de Krom MC, Knipschild PG, Kester AD, Thijs CT, Boekkooi PF, Spaans F (1992) Carpal tunnel syndrome: prevalence in the general population. J Clin Epidemiol 45:373–376
3. Ferry S, Pritchard T, Keenan J, Croft P, Silman AJ (1998) Estimating the prevalence of delayed median nerve conduction in the general population. Br J Rheumatol 37:630–635
4. Becker J, Nora DB, Gomes I, Stringari FF, Seitensus R, Panosso JS et al (2002) An evaluation of gender, obesity, age and diabetes mellitus as risk factors for carpal tunnel syndrome. Clin Neurophysiol 113:1429–1434
5. Lam N, Thurston A (1998) Association of obesity, gender, age and occupation with carpal tunnel syndrome. Aust N Z J Surg 68:190–193
6. Pal B, Morris J, Keenan J, Mangion P (1997) Management of idiopathic carpal tunnel syndrome (ICTS): a survey of rheumatologists' practice and proposed guidelines. Br J Rheumatol 36:1328–1330
7. Palmer KT, Harris EC, Coggon D (2007) Carpal tunnel syndrome and its relation to occupation: a systematic literature review. Occup Med (Lond) 57:57–66
8. Alcalde M, D'Agostino MA, Bruyn GA, Moller I, Iagnocco A, Wakefield RJ, Force OUT et al (2012) A systematic literature review of US definitions, scoring systems and validity according to the OMERACT filter for tendon lesion in RA and other inflammatory joint diseases. Rheumatology (Oxford) 51:1246–1260
9. Backhaus M, Burmester GR, Gerber T, Grassi W, Machold KP, Swen WA, Working Group for Musculoskeletal Ultrasound in the ESCoICSiTT et al (2001) Guidelines for musculoskeletal ultrasound in rheumatology. Ann Rheum Dis 60:641–649
10. Naredo E, Moller I, Moragues C, de Agustin JJ, Scheel AK, Grassi W et al (2006) Interobserver reliability in musculoskeletal ultrasonography: results from a "Teach the Teachers" rheumatologist course. Ann Rheum Dis 65:14–19
11. Nakamichi K, Tachibana S (1998) Histology of the transverse carpal ligament and flexor tenosynovium in idiopathic carpal tunnel syndrome. J Hand Surg [Am] 23:1015–1024
12. Uchiyama S, Itsubo T, Nakamura K, Kato H, Yasutomi T, Momose T (2010) Current concepts of carpal tunnel syndrome: pathophysiology, treatment, and evaluation. J Orthop Sci 15:1–13
13. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostino MA et al (2005) Musculoskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol 32: 2485–2487
14. Szkudlarek M, Court-Payen M, Jacobsen S, Klarlund M, Thomsen HS, Ostergaard M (2003) Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. Arthritis Rheum 48:955–962
15. Ettema AM, Amadio PC, Zhao C, Wold LE, An KN (2004) A histological and immunohistochemical study of the subsynovial connective tissue in idiopathic carpal tunnel syndrome. J Bone Joint Surg Am 86-A:1458–1466
16. Neurologie NVv (2006) Richtlijn: Diagnostiek en behandeling van het carpale-tunnelsyndroom

3

Very different performance of the power Doppler modalities of several ultrasound machines ascertained by a microvessel flow phantom.

Ten Cate DF, Luime JJ, van der Ven M, Hazes JM, Kooiman K, de Jong N, Bosch JG

ABSTRACT

Introduction: In many patients with rheumatoid arthritis (RA) subclinical disease activity can be detected with ultrasound (US), especially using power Doppler US (PDUS). However, PDUS may be highly dependent on the type of machine. This could create problems both in clinical trials and in daily clinical practice. To clarify how the PDUS signal differs between machines we created a microvessel flow phantom.

Methods: The flow phantom contained three microvessels (150, 1000, 2000 microns). A syringe pump was used to generate flows. Five US machines were used. Settings were optimised to assess the lowest detectable flow for each US machine.

Results: The minimal detectable flow velocities showed very large differences between the machines. Only two of the machines may be able to detect the very low flows in the capillaries of inflamed joints. There was no clear relation with price. One of the lower-end machines actually performed best in all three vessel sizes.

Conclusions: We created a flow phantom to test the sensitivity of US machines to very low flows in small vessels. The sensitivity of the power Doppler modalities of 5 different machines was very different. The differences found between the machines are probably caused by fundamental differences in processing of the PD signal or internal settings inaccessible to users. Machines considered for PDUS assessment of RA patients should be tested using a flow phantom similar to ours. Within studies, only a single machine type should be used.

INTRODUCTION

Rheumatoid arthritis (RA) is a common disease with a prevalence of around 1% worldwide (1). RA is potentially an invalidating disease (2), but early diagnosis in the so-called window of opportunity (3,4) and treatment according to a treat-to-target protocol (5) can optimise the outcome for RA patients. Adding ultrasound (US) to the diagnostic workup and monitoring of treatment may provide even better results. In rheumatological US both greyscale and power Doppler (PD) are used, of which PDUS seems to have the largest value. PDUS has the potential to reclassify patients to a higher joint group according to the 2010 classification criteria for RA, increasing the risk for undifferentiated arthritis to be definite RA (6). Furthermore, the presence of PDUS inflammation in joints that are not swollen at clinical examination has shown to be clinically relevant in patients in remission of RA, since it predicts occurrence of flare and erosive progression (7-10). Correct assessment of the presence and absence of a PD signal indicating the presence of inflammation is therefore vital in rheumatological US.

4
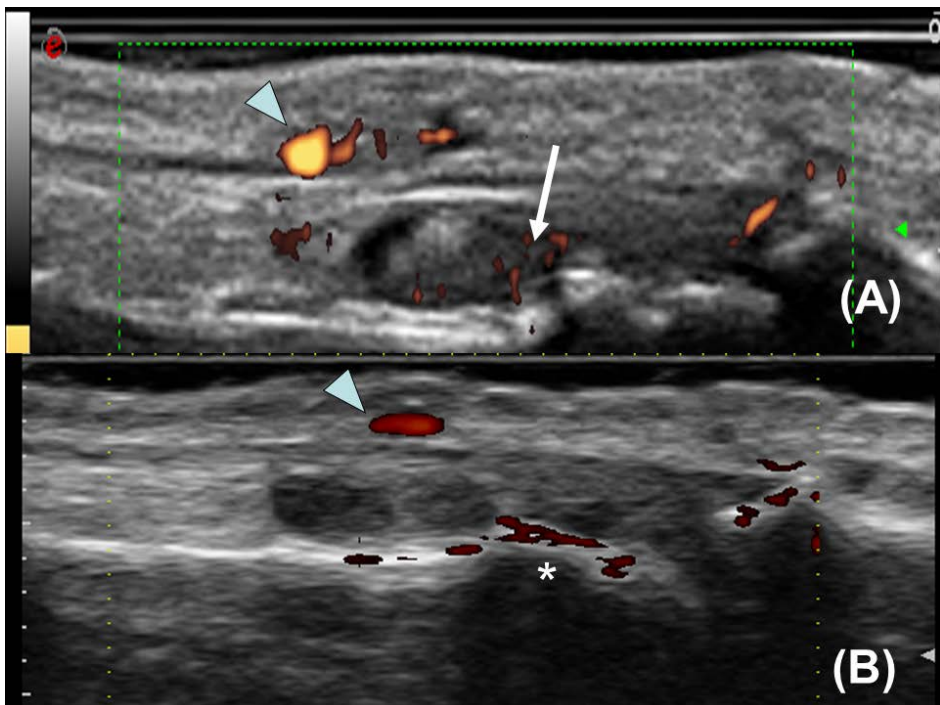
Figure 1. Different performance of the power Doppler ultrasound modality of two machines in one patient. (A) Machine B, presence of a positive power Doppler signal (arrow) within the region of synovial proliferation. (B) Machine A, this signal is absent. Arrowhead, vessel; *noise on cortical surface.

However, PDUS may be highly dependent on the type of US machine used (11,12). We also observed this in our centre (see Figure 1). If there are indeed large differences in the performance of PDUS per machine, the choice of machine might be essential for a valid detection of inflammation. Using different machines within a multi-machine study or during patient treatment could then have a detrimental impact on treatment decisions or study outcome.

To quantify the suspected differences in PD sensitivity of different machines in an objective way, we decided to perform an *in vitro* experiment. To compare the PD function of different US machines one could use a flow phantom. This flow phantom should mimic the tissue that is scanned by PDUS in rheumatology; that is, very small vessels and very low flows. To our knowledge, no studies have been conducted investigating the size of capillaries and the blood flow velocity in an inflamed joint, but there are data on capillaries in healthy subjects' nail folds and capillaries in periulcerous regions. These capillaries have a diameter of around 30 μm and the blood flow velocity can be as low as 0.5 mm/second (13,14). Flow phantoms previously presented did not compare US machines (15), used vessels that were considerably larger than capillaries, or assessed many capillaries close to each other at once, making it impossible to evaluate the flow velocity in the individual vessels (11,12,16,17). For these reasons, we created a new flow phantom with a very small, single vessel to obtain the lowest detectable flow velocity of five US machines. Two additional larger vessels were included in the phantom for comparison with the literature (11,12).

METHODS

Phantom

The flow phantom (Figure 2) consisted of an acrylic (polymethyl methacrylate) container filled with tissue mimicking material, according to a previously published recipe (18). In this tissue-mimicking material we placed three microvessels: 150 μm (inner diameter) vessel made of polyethylene terephthalate glycol-modified (Paradigm Optics, Vancouver, WA, USA), and 2,000 μm and 1,000 μm (inner diameter) vessels made of silicone (Eriks bv, Alkmaar, the Netherlands). These two vessels were included to compare our phantom with already published studies (11,12). Initially we used vessels with diameters of 50 μm and 100 μm also made of polyethylene terephthalate glycol-modified, but these were blocked almost instantly. The blood-mimicking fluid (BMF) was based on the recipe by Ramnarine and colleagues (19). Briefly, 91.07% (w/w) demineralised water, 1.18% (w/w) dextran (average 150 kDa, D4876; Sigma-Aldrich,
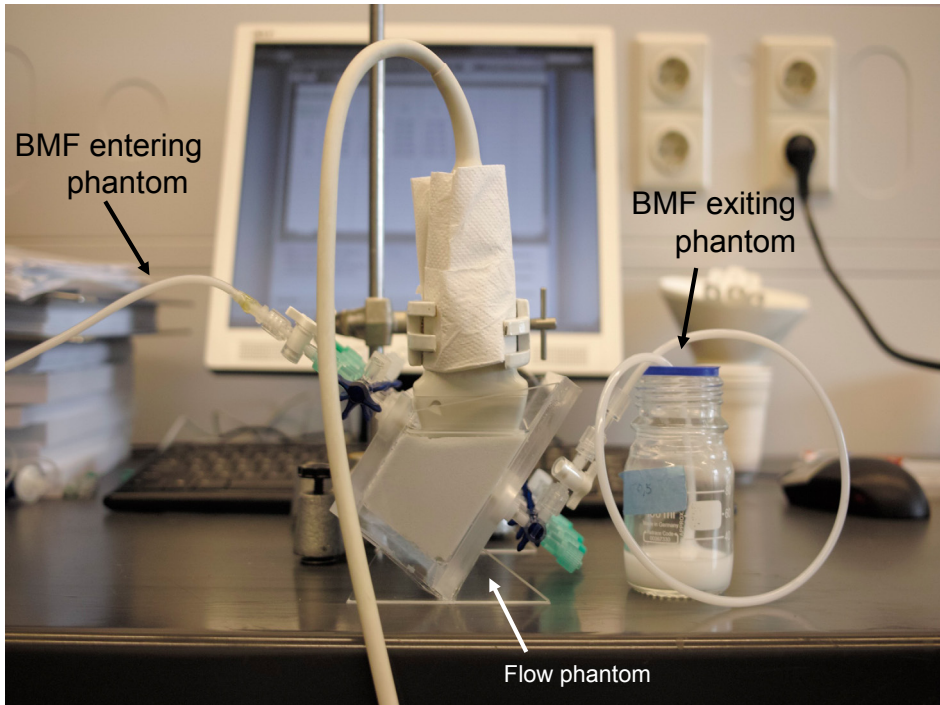
Figure 2. Flow phantom with a fixated probe. BMF, blood-mimicking fluid.

Zwijndrecht, the Netherlands), 0.90% (w/w) ICI supersonic N surfactant, 5.03% (w/w) glycerol, and 1.82% w/w orgasol particles (5 μm in diameter; Arkema B.V., Rotterdam, The Netherlands) were mixed using a magnetic stirrer. The BMF was then filtered using a 40 μm sieve (352340; BD, Breda, the Netherlands) and degassed using a vacuum pump. Compared with the original recipe by Ramnarine and colleagues, our BMF contained half the amount of dextran and glycerol – this made our BMF less viscous, which was necessary to prevent blockage of the vessels. A syringe pump (Hugo Sachs Elektronik, MarchHugstetten, Germany) was used to generate flows. This pump can produce regular flows as low as 1.28 pl/minute. For each vessel size, flow settings (ml/hour) were calculated that corresponded to average flow velocities ranging from 40 to 0.005 mm/second, using the following equation, where Q is flow (m3/second), *Vavg* is the average flow velocity (m/second) and R is the radius (m):

$$Q = V\,avg \times \pi R2$$

The actual volume transported through the vessels was tested by turning on the pump, and completely filling the vessel until drops of BMF came out of the capillary. A complete number of drops were captured in a container while recording the time.

This container was weighed before and after this experiment on a microbalance. With the relative density of the BMF we calculated the flow (transported volume per time).

Experiment

The lowest detectable flow for each machine/vessel diameter combination was defined as the flow that still resulted in a continuous PDUS signal (Figure 3). First the pump was set to a high flow, and then gradually decreased in steps until a continuous PD signal could just still be detected. The value of the lowest flow was recorded. Between each change of pump flow we waited 5 minutes for the system to reach stable flow velocities. For each lowest detectable flow per vessel we stored an image and recorded the machine settings used to acquire this image.

Figure 3. Continuous PDUS signal in a 1 mm vessel on Machine B.

Ultrasound machines and settings

Five available US machines were tested (Table 1). Machines A and B are used in our Department of Rheumatology in daily clinical practice. Machine C is a high-end machine for general imaging. Machines D and E are specialised research machines, the latter a highly specialised machine for high-frequency small animal imaging. Four machines operated at or around the most common frequency of 10 MHz (Machines A, B, D and E), and one (Machine C) operated lower at a frequency between 3 and 9 MHz (actual frequency not displayed on this machine). Settings on all machines were optimised to detect lowest flows by adjusting pulse repetition frequency (PRF)/velocity range, wall filters, Doppler frequency and Doppler gain. In general this meant for all vessels using the lowest wall filter, the lowest velocity range or PRF and the highest suitable Doppler gain with respect to noise level. One experienced musculoskeletal ultrasonographer (DFTC) performed all US examinations.

Table 1. Machines tested in alphabetical order and probes used

Machine A	Aloka α7 (probe UST-5411)
Machine B	Esaote MyLab60 (probe LA 435)
Machine C	Philips iU22 (probe L9-3)
Machine D	Ultrasonix SonixTouch (probe L14-5/38)
Machine E	Visualsonics Vevo2100 (probe MS200)

RESULTS

We found that the pump was accurate enough for our purposes, especially when taking into account the very low flows used (see Table 2). The lowest detectable flow velocities in the different vessels are presented in Table 3. These differed very much, by a factor of 100 between machines. This was the case for all vessel sizes. In the smallest vessel (150 μm), which most resembles the situation in an inflamed joint, two machines (Machines D and E) could not detect a positive PD signal at all at any flow velocity. For the other vessels, the minimal detectable velocity ranged from 0.11 mm/second (Machine B) to 11.1 mm/second (Machine A). The settings were optimised for the detection of lowest flow. See Table 4 for the settings of PRF/velocity range, wall filter and Doppler frequency per machine for the smallest vessel.

Table 2. Measuring the reliability of the pump.

Vessel size(micron) / Flow (ml/h)	2000	1000	150	150	150
Set flow	11.31	3.142	0.141	0.070	0.035
Measured flow	10.68	2.948	0.276	0.108	0.049

ml/h: millilitre per hour

Table 3: Lowest detected flow velocity (in mm/s), still resulting in a continuous positive PDUS signal

	Flow velocity (mm/s)		
Vessel size (micron) / Machine	2000	1000	150
A	4	2.22	11.1
B	0.005	0.06	0.11
C	1	0.56	1.68
D	1	0.56	N.D.
E	0.5	0.33	N.D.

mm/s: millimetre per second; N.D. = none detected.

Table 4: Settings for detection of lowest flow velocity in the 150 micron vessel

Machine	PRF / Velocity range	Wall filter	Doppler frequency
A	1.3 cm/s	Level 1	8 MHz
B	125 Hz	Level 1	10 MHz
C	150 Hz	15 Hz	R1 (Actual frequency not displayed on this machine)
D (no flow detected in this vessel)	200 Hz	Level 1	10 MHz
E (no flow detected in this vessel)	1000Hz	Low	12.5 MHz

DISCUSSION

We showed that the sensitivity of the PD modalities of five US machines (three machines used in the clinic and two used for research) was very different, using a microvessel flow phantom. The very large differences found between the machines are only partly explained by each machine's Doppler frequency, lower limits of PRF and

wall filter settings, but are most probably caused mainly by fundamental differences in processing the PD signal or internal settings inaccessible to users. There was no clear relation with price or technical sophistication of the machines: a lower-end machine (Machine B) performed best for all three vessels, while mid-range and high-end research machines (Machines D and E) did not detect any flow in the smallest vessel, against expectations.

Only one machine of the five (Machine B) could detect the low flow velocity in capillaries that based on previous research are estimated to be between 0.5 and 1 mm/second. Machine C came close to this limit, which underlines our conclusion that the observed differences are mainly caused by differences in processing of the signal, since the probe that was available for Machine C had a bandwidth of only 3 to 9 MHz. When a high-frequency probe would have been used with this machine, it might also have been able to detect flow less than 1 mm/second in the smallest vessel. The other machines did not perform appropriately according to this limit.

As mentioned above, flow phantoms have been published in the literature before (11,12,15-17). However, when comparing the PD modalities of different US machines it is essential to use small, individual vessels. A positive PD signal depends on the total detected Doppler signal power within the range gate (the colour Doppler pixel size, typically <1 mm). This power depends on the number of particles that have a velocity above a certain threshold. This threshold is determined by the wall filter, the PRF/velocity range. Whether a PD signal is actually detected/displayed is also dependent on the noise level of the system and the system's ability to suppress clutter and signal from stationary targets. If the vessel diameter is larger than the gate size, the velocity threshold will determine the lowest detectable velocity. This explains why the minimum velocities found for 1 and 2 mm vessels are similar.

However, if the velocity is the same but the vessel is much smaller than the gate size, the number of moving particles will be lower and more stationary tissue will be inside the gate range. The tissue suppression and noise level then become more important and the minimal detectable velocity will be raised. This means that a phantom with a vessel too large in diameter (11,12) may use a flow velocity similar to that in vessels in an inflamed joint, but more particles are inside one pixel in the phantom situation (*in vitro*) compared with the situation in an inflamed joint (*in vivo*). This can possibly cause a positive PD signal based on the large number of particles. In a flow phantom using a bundle of capillaries (16,17) one can never know for sure the flow velocity in each vessel. The possibility therefore remains that the flow is very high in a few capillaries, causing a positive PD signal solely based on the high flow velocity of particles in these few capillaries.

A study comparing Machines A and B (older versions than in our study) on a 1,000 μm flow phantom has been published in the past (12). These older versions of the

machines were ranked regarding sensitivity the same as in our study. However, in our study Machine B detected a considerably lower flow compared with the previous study: 0.06 mm/second in our study versus

1.3 mm/second in the previous. Machine A detected a twofold lower flow in our study: 2.2 mm/second in our study versus 3.9 mm/second in the earlier study.

Another study tested an earlier, single-element version of Machine E (VisualSonics, Toronto, Canada Vevo 770;) on a microvessel flow phantom with vessel dimensions similar to ours (160 µm) (15). In this microvessel, the Vevo 770 did detect flows as low as 0.5 mm/second. In our study Machine E did not detect any flow in the smallest vessel (150 µm). A possible explanation for this higher sensitivity for low flows could be that the Vevo 770 uses a mechanically steered probe with a single element, as opposed to the array probe we used on Machine E in our study. In general, the Doppler processing of a single-element system can be very different from that of an array system.

Some observations raised discussion within our research group. One of these discussions was about the very low flows detected by Machine B in the 2,000 µm vessel. To verify this finding the experiment was repeated several times by two observers (DFTC and MvdV), which resulted in similar findings. When setting the flow slightly lower, the signal disappeared. Therefore, we think the measured flow is correct. A possible explanation for this low limit is that the PRF can be set to a very low level and the wall filter cutoff frequency is probably also very low, in combination with a good clutter suppression. However, in *in vivo* situations, normal tissue or probe motion will prevent detecting such extremely low flows.

Another observation that raised discussion is the lower flow detected in the 1,000 µm vessel as compared with the 2,000 µm vessel. A reason for this could be that the flow velocity profile in the 1,000 µm vessel is shaped differently, as compared with the 2,000 µm vessel, resulting in a larger difference between average flow and maximum flow. This may even have been reinforced by compression of the smaller vessel by the tissue-mimicking material. This means the average flow velocity is actually higher than estimated, since the calculation is quadratically dependent on the microvessel diameter. If the maximum velocity of the peak flow is slightly higher than the wall filter cutoff, this results in a positive PD signal. The peak flow may therefore be rather similar in the 2,000 µm and the 1,000 µm vessels, but due to the shape of the flow profile this corresponds to a lower average flow velocity in the 1,000 µm vessel. While the true value for the flow velocities may differ from the calculated values, this difference is the same for all machines, so the comparison between machines is still valid per vessel.

A drawback of our study is that we have made assumptions on the capillary sizes and flow velocities in inflamed joints based on papers published on healthy subjects

and periulcerous regions. This may not be entirely correct. Therefore, at present it is crucial to ascertain the flow velocities and capillary sizes in inflamed joints. With this information the minimal flows that rheumatological US machines need to be able to detect will be known.

Nonetheless, for a reliable and reproducible detection of very low flows in inflamed joints, the choice of the US machine and its settings seems very important. Caution should be exercised when conducting a multi-machine trial or when making treatment decisions based on PDUS. Our flow phantom could be used to decide which US machine to use both in clinical practice and in clinical trials.

4

CONCLUSIONS

We created a flow phantom to test the sensitivity of US machines to very low flows in small vessels. We found that the sensitivity of the PD modalities was very different between five US machines. Based on the results of our study it would be advisable to standardise and validate US machines both for rheumatological clinical practice and for clinical trials. Our phantom could be used for this purpose.

ACKNOWLEDGEMENTS

Note: Abbreviations are explained in the abbreviations list in the appendix.

REFERENCES

1. Silman AJ, Pearson JE: Epidemiology and genetics of rheumatoid arthritis.*Arthritis Res* 2002, 4: S265–S272.

2. Gordon P, West J, Jones H, Gibson T: A 10 year prospective followup of patients with rheumatoid arthritis 1986–96. *J Rheumatol* 2001, 28:2409–2415.

3. Mottonen T, Hannonen P, Korpela M, Nissila M, Kautiainen H, Ilonen J, Laasonen L, Kaipiainen-Seppanen O, Franzen P, Helve T, Koski J, Gripenberg-Gahmberg M, Myllykangas-Luosujärvi R, Leirisalo-Repo M,FIN-RACo Trial Group, FINnish Rheumatoid Arthritis Combination therapy: Delay to institution of therapy and induction of remission using single-drug or combination-disease-modifying antirheumatic drug therapy in early rheumatoid arthritis. *Arthritis Rheum* 2002, 46: 894–898.

4. Quinn MA, Emery P: Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention.*Clin Exp Rheumatol* 2003, 21:S154–S157.

5. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, Kincaid W, Porter D: Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004, 364:263–269.

6. Filer A, De Pablo P, Allen G, Nightingale P, Jordan A, Jobanputra P, Bowman S, Buckley CD, Raza K: Utility of ultrasound joint counts in the prediction of rheumatoid arthritis in patients with very early synovitis. *Ann Rheum Dis* 2011, 70:500–507.

7. Brown AK, Conaghan PG, Karim Z, Quinn MA, Ikeda K, Peterfy CG, Hensor E, Wakefield RJ, O'Connor PJ, Emery P: An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. *Arthritis Rheum* 2008, 58:2958–2967.

8. Foltz V, Gandjbakhch F, Etchepare F, Rosenberg C, Tanguy ML, Rozenberg S, Bourgeois P, Fautrel B: 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.

9. Peluso G, Michelutti A, Bosello S, Gremese E, Tolusso B, Ferraccioli G: Clinical and ultrasonographic remission determines different chances of relapse in early and long standing rheumatoid arthritis. *Ann Rheum Dis* 2011, 70:172–175.

10. 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, 48:1092–1097.

11. Browne JE, Watson AJ, Hoskins PR, Elliott AT: Validation of a sensitivity performance index test protocol and evaluation of colour Doppler sensitivity for a range of ultrasound scanners. *Ultrasound Med Biol* 2004, 30:1475–1483.

12. Koski JM, Saarakkala S, Helle M, Hakulinen U,Heikkinen JO, Hermunen H: Power Doppler ultrasonography and synovitis: correlating ultrasound imaging with histopathological findings and evaluating the performance of ultrasound equipments. *Ann Rheum Dis* 2006, 65:1590–1595.

13. Stucker M, Baier V, Reuther T, Hoffmann K, Kellam K, Altmeyer P: Capillary blood cell velocity in human skin capillaries located perpendicularly to the skin surface: measured by a new laser Doppler anemometer. *Microvasc Res* 1996, 52:188–192.

14. Stucker M, Huntermann C, Bechara FG, Hoffmann K, Altmeyer P: Capillary blood cell velocity in periulcerous regions of the lower leg measured by laser Doppler anemometry. *Skin Res Technol* 2004, 10:174–177.

15. Pinter SZ, Lacefield JC: Detectability of small blood vessels with high-frequency power Doppler and selection of wall filter cut-off velocity for microvascular imaging. *Ultrasound Med Biol* 2009, 35: 1217–1228.

16. Kamishima T, Nishida M, Horie T, Narita A, Sagawa A, Henmi M, Shirato H, Terae S: Power Doppler signal calibration using capillary phantom for pannus vascularity in rheumatoid finger joint. A pilot study. *Clin Exp Rheumatol* 2011, 29:1057.

17. Veltmann C, Lohmaier S, Schlosser T, Shai S, Ehlgen A, Pohl C, Becher H, Tiemann K: On the design of a capillary flow phantom for the evaluation of ultrasound contrast agents at very low flow velocities. *Ultrasound Med Biol* 2002, 28:625–634.

18. Teirlinck CJ, Bezemer RA, Kollmann C, Lubbers J, Hoskins PR, Ramnarine KV, Fish P, Fredeldt KE, Schaarschmidt UG: Development of an example flow test object and comparison of five of these test objects, constructed in various laboratories. *Ultrasonics* 1998, 36:653–660.

19. Ramnarine KV, Nassiri DK, Hoskins PR, Lubbers J: Validation of a new blood-mimicking fluid for use in Doppler flow test objects. *Ultrasound Med Biol* 1998, 24:451–459.

4
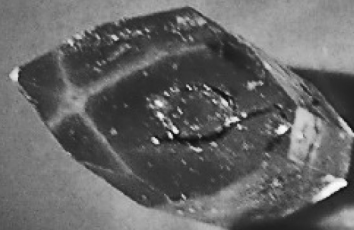
Does the intraclass correlation coefficient always reliably express reliability?

Ten Cate DF, Luime JJ, Hazes JM, Jacobs JW, Landewé R

The reliability of ultrasonographic examination in rheumatology is a matter of ongoing debate. In an article published in Arthritis Care & Research, Cheung et al systematically reviewed the reliability of B-mode and power Doppler (PD) ultrasonography (US) to detect synovitis in rheumatoid arthritis (RA) in 35 studies, comprising 1,415 patients and reporting high interobserver and intraobserver reading reliability, especially of PD (1). However, the "reliability of the measures of reliability" is not beyond question.

Among other measures of reliability, the intraclass correlation coefficient (ICC) had been used to assess reliability. Two challenges exist in the use and interpretation of the ICC. First, the ICC is highly dependent on the heterogeneity of the study sample, and as a consequence is generalizable only to samples with a similar variation (2). The ICC is basically a signal-to-noise ratio. This may be difficult to comprehend conceptually, but it can be clarified by looking at the following equation:

$$ICC = \frac{\text{Variance (patients)}}{\text{Variance (patients)} + \text{Variance (observer)} + \text{Variance (error)}}$$

5

This equation states that the heterogeneity of patients under investigation determines for a large part the value of the ICC. When the variance between patients, shown as Variance (patients), is low, the ICC is likely to be low as well, and vice versa. This also applies to rheumatologic US. When only a few of the total number of joints under investigation show signs of synovitis, which means a low variance between patients, the ICC will probably also be low and largely independent of the level of variance between observers. So the outcome of the ICC does not necessarily express the reliability reliably, nor the variance between observers and variance due to error.

The second issue regarding the ICC is that its value depends on which ICC equation is used. For different study designs there are several different ICC equations (3,4). The most important distinction to be made is between the ICC for agreement and the ICC for consistency, which consists of different equations (2,4). It is possible that observers fully agree on ranking patients into a low, intermediate, or high level of pathology according to the assessed scores, resulting in a high ICC for consistency among observers. However, this does not necessarily mean that the observers also reach agreement in the raw values of the scores given, which is the basis of the ICC for agreement. This is illustrated in Table 1. Two observers are asked to score the disease activity of 3 patients on a scale from 0 –100. This situation would give an "almost perfect reliability" when calculating the ICC for consistency, since the 3 patients are ranked according to disease activity scores in the same order by both observers. However, when calculating the ICC for agreement, the result will be "poor reliability," since the scores between the observers clearly differ by ~ 10-fold.

Table 1. The disease activity of 3 patients rated by 2 observers

	Observer A	Observer B
Patient 1	3	59
Patient 2	6	75
Patient 3	9	93

In studies using the ICC, the extent of heterogeneity within the study population should be analyzed and described, as heterogeneity clearly influences the ICC. An example of how this can be done is described in an Outcome Measures in Rheumatology Clinical Trials article on magnetic resonance imaging (5). Furthermore, authors using an ICC should describe which method is used to calculate it and the rationale for their choice. This way, readers can better appreciate the reported reliability. In the articles reviewed by Cheung et al, only some make reference to which formula of ICC has been used (6), and a measure of heterogeneity, as a variable in the outcome of the ICC, is not given in any of these articles. These issues are the axe at the root of the robustness of the review of Cheung and colleagues, and should, in our opinion, have been acknowledged and discussed by Cheung et al in their article.

Note: Abbreviations are explained in the abbreviations list in the appendix.

REFERENCES

1. Cheung P, Dougados M, Gossec L. Reliability of ultrasonography to detect synovitis in rheumatoid arthritis: a systematic literature review of 35 studies (1,415 patients). Arthritis Care Res (Hoboken) 2010;62:323–34.
2. De Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol 2006; 59:1033–9.
3. Shrout P, Fleiss J. Intraclass correlations: uses in assessing rater reliability. Psych Bull 1979;86: 420 – 8.
4. McGraw K, Wong S. Forming inferences about some intraclass correlation coefficients. Psych Meth 1996;1:30 – 46.
5. Lukas C, Braun J, van der Heijde D, Hermann KG, Rudwaleit M, Ostergaard M, et al. Scoring inflammatory activity of the spine by magnetic resonance imaging in ankylosing spondylitis: a multireader experiment. J Rheumatol 2007;34: 862–70.
6. Naredo E, Collado P, Cruz A, Palop MJ, Cabero F, Richi P, et al. Longitudinal power Doppler ultraso-nography assessment of joint inflammatory activity in early rheumatoid arthritis: predictive value in disease activity and radiologic progression. Arthritis Rheum 2007;57:116 –24.

5

CHAPTER 5.2

The most reliable probe position in the ultrasonographic examination of the wrist in rheumatoid arthritis.

Dopazo Gonzalez N, Ten Cate DF, Swen WA, Mera Varela A,
Insua Vilarino SA, Perez-Pampin E, Gomez-Reino JJ, Luime JJ

ABSTRACT

Objective: This study aims to evaluate the inter-observer reliability of the ultrasonographic examination of the wrist in RA patients between 3 examiners and 3 probe positions.

Methods: Fifty-three RA patients were recruited at the University Clinical Hospital of Santiago de Compostela in Spain for ultrasonographic examination of the wrist. Ultrasonography (US) was performed on both wrists using a GE LOGIQ 9 machine, using three probe positions: Lister's Tubercle to digit II (position 1), Lister's Tubercle to digit III (position 2) and ulnocarpal (position 3), from the anatomic medial orientation. Three examiners (2 experienced ultrasonographers and 1 junior ultrasonographer) scored synovitis according to a 0–3 semiquantitative scoring system. Inter-observer reliability was expressed using the ICC (A,1).

Results: For grey-scale ultrasound (GSUS) the inter-observer reliability (ICC(A,1)) (single measure, agreement definition) ranged from 0.35 for the ulnocarpal joint, position 3, to 0.60 in both position 1 for the radiocarpal joint and position 2 for the inter-carpal joint. Using power Doppler ultrasound (PDUS) the inter-observer reliability (ICC(A,1)) ranged from 0.36 in position 3, to 0.52 both in position 1 and 2 for the radiocarpal joint.

Conclusion: The reliability of the GSUS-examination of the wrist joints of RA patients with GSUS shows highest, moderate reliability using the anatomical landmarks Tubercle of Lister and digit III (position 2). The reliability of the PDUS examination was similar and moderate in both position 1 (Lister's Tubercle to digit II) and position 2 (Lister's Tubercle to digit III). The reliability was poorest for position 3 (the anatomic medial view of the ulnocarpal joint) in both the GSUS and PDUS examination. This study suggests that position 2 should be used in clinical trials and daily practice.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic and progressive inflammatory disorder primarily affecting synovium, cartilage and bone. Early intervention and tight disease control is crucial, improving the prognosis of RA substantially (1-2). Further improvement might be expected by early detection of synovitis and erosions, using modern imaging techniques such as ultrasonography (US) (3-4) and magnetic resonance imaging (MRI) (5).

US is a growing imaging modality with great potential as an outcome measure in RA. But at the same time, US is perceived as imperfect and operator dependent (3-6). The operator dependency and the lack of clear criteria for interpretation of US images cause challenges in the interpretation and acquisition of US images. The wrist is one of the most difficult joints of the body to assess ultrasonographically (7). For this joint, MRI outperforms US (8-9). The complex anatomy of the wrist, the layer of soft tissue between the probe and the joint, minimal tissue contrast and lack of a reproducible window for visualising these structures complicates imaging of the wrist (7). Because of these difficulties, reliability is an issue. However previous studies have usually shown reliability data on a combination of joints and not for the wrist only (10). They also did not describe distinct landmarks for probe positioning, leaving the exact location where pathology was evaluated open for discussion. These issues, in combination with a lack of consensus of probe positioning, has led to little and equivocal data on intra- and interobserver reliability of US scanning of the wrist (7, 10-11). This issue is recognised by the OMERACT (4).

Using distinct anatomical landmarks may improve reliable probe positioning, providing more consistent results between rheumatologists when evaluating the wrist. We therefore evaluated the inter-observer reliability of 3 different ultrasonographic probe positions of the dorsal and anatomic medial side of the wrist of RA patients using Grey-scale Ultrasound (GSUS) and Power Doppler Ultrasound (PDUS) by 3 examiners.

PATIENTS AND METHODS

Study design

A cross-sectional study, approved by the local ethics committee, was conducted at the University Clinical Hospital of Santiago de Compostela, Spain. Patients were invited if diagnosed with RA-ACR1987, aged over 18 years and attending the outpatient rheumatology clinic, or day clinic between May 2010 and July 2010. Subjects' written

informed consent was obtained according to the Declaration of Helsinki before participation. Exclusion criteria were a history of fractured wrist bones or ulna/radius, an intra-articular injection with glucocorticosteroids in the wrist <6 months ago and/or malformation of the wrist inhibiting correct probe positioning. No requirements were made regarding the presence of clinical wrist synovitis.

US examination

Both wrists of the patients were scanned from the dorsal side using three probe positions. The positions were chosen because they are defined by anatomically distinct landmarks (Fig. 1a).

- Position 1: From Lister's Tubercle to the proximal end of the second metacarpal bone (DIG II); imaging the radiocarpal joint (RCJ) and intercarpal joint (ICJ).
- Position 2: From Lister's Tubercle to the proximal end of the third metacarpal bone (DIG III); imaging the RCJ and the ICJ.
- Position 3: From the anatomic medial side of the ulnar styloid process to the region between the proximal end of the fourth (DIG IV) and fifth metacarpal bone (DIG V); imaging the ulnocarpal joint (UCJ). (Note that the position of the hand should be different for position 3 than presented in the figure))

The US scans were made using a GE LOGIQ 9 machine with a 5–13MHz multifrequency linear array transducer. Patients were set in front of the examiner, with the palm of the hand placed flat on the table in a neutral position for the first two positions. For the third position the hand was placed in maximum abduction, seen from the anatomical position, ensuring good probe positioning. All patients were examined in the same room with a constant temperature. Evaluation of one patient took ~10 minutes including documentation.

Synovitis was scored real-time, using a semiquantitative scoring system (grade 0=absence, grade 1=mild, grade 2=moderate, and grade 3=severe synovitis), using a horizontal line from the upper surface of the distal radius parallel to the upper side of US screen (Fig. 1b). For position 3 this line ran from the upper surface of the ulna to the upper surface of the triquetrum (Fig. 1c). A distension of the joint capsule beneath this horizontal line was graded 1, at or above this line was graded 2 and bulging above the line was graded 3. For Power Doppler (PD) we used the following settings: colour gain at disappearance of colour noise, Pulse Repetition Frequency (PRF) at 0.5 kHz, frequency at 7.5 MHz and wall filter (WF) at 83 Hz. We adjusted the size and position of the colour box to include the subcutaneous tissue to recognise artefacts caused by vessels above the joint. The PD signal was evaluated using the semiquantitative PD scoring system of Szkudlarek (3).
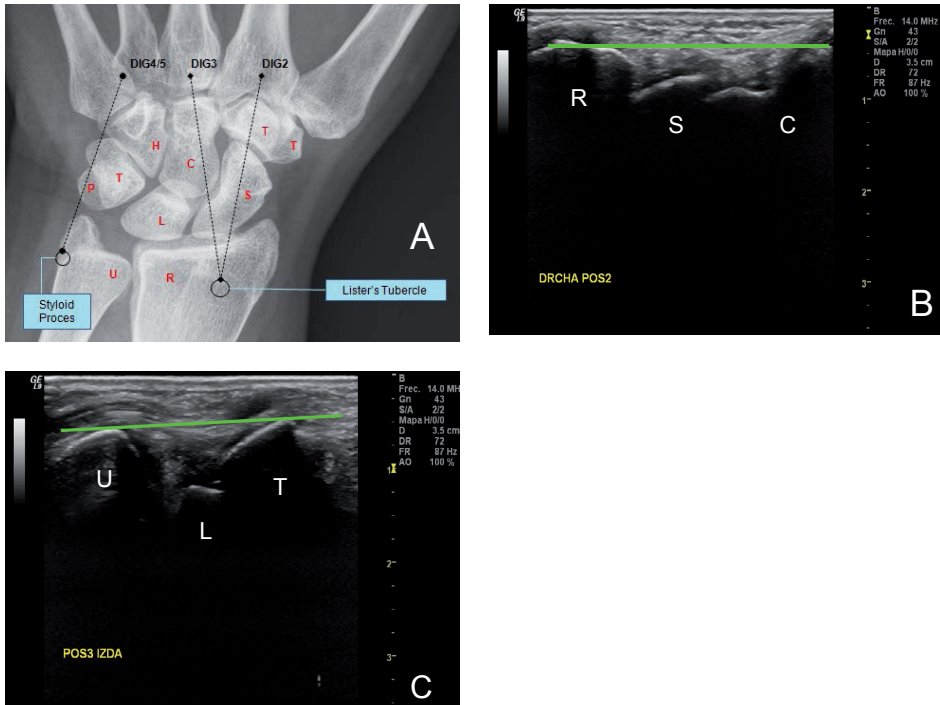
Figure 1. A. Radiography of the wrist, with the 3 probe positions. R= Radius, U= Ulna, S= Scaphoid, L=Lunatum, T= Triquetrum (left), P= Pisiforme, H=Hamatum, C= Capitatum, T= Trapezoideum (middle),T= Trapezium (right).
B. Horizontal line from the surface of the radius, parallel to upper side of the US screen. Used for evaluation in the first and second position. (US image in position 2: R= Radius, S= Scaphoid, C=Capitatum.)
C. Line from upper surface of the ulna to upper surface of the triquetrum. Used for third position. (US image in position 3: U=Ulna, L=Lunatum, T= Triquetrum.)

Ultrasonographers

Each patient was evaluated by two rheumatologists experienced in musculoskeletal US and one junior doctor trained for this occasion (>70 supervised US-wrist examinations prior to this study). The scans were performed independently, and blinded for clinical examination. Consensus regarding the scoring protocol was reached prior to the study and documentation was standardised per patient.

Additional data collection

Data on age, gender, disease duration, rheumatoid factor (RF), anticyclic citrullinated peptide (anti-CCP), Sharp-van der Heijde score, HAQ-score (12), DAS28 and current treatment were collected after US examination. Data on C-reactive protein (CRP) concentration and erythrocyte sedimentation rate (ESR) were obtained using each patient's most recent laboratory results. An independent rheumatologist, blinded to US findings, obtained a DAS28 (13).

Statistical analysis

The sample size was calculated according to Streiner (14). With an expected ICC(A,1) of 0.8, a confidence interval of 0.1 and 3 observations per patient, 46 patients were needed. Descriptive statistics were performed for patient and disease characteristics. Reliability was calculated by the intraclass correlation coefficient (ICC (A,1)) (15), single measure, agreement definition in R (version 2.7.1). The inter-observer reliability ICC (A,1) was determined per position (1, 2, or 3) and per joint (radiocarpal (RC) and intercarpal (IC)). The scores for the right and left wrist were combined. This study was conducted following the STROBE statement for observational studies (16).

RESULTS

Patient characteristics (Table 1)

The 53 RA patients had a median disease duration of 57 months [range: 0– 354 months], a heterogeneous disease activity (median DAS28 of 2.83 [range 0.19–6.41]) and were reasonably selfsufficient (median HAQ of 0.88 [range: 0–2.8]). DAS28 indicated remission in 17 patients (32.1%). Four patients (7.5%) had severe disease activity. Clinically, 32 wrists were tender, of which 24 were also swollen. Four additional wrists were only swollen. 38 patients (71.6%) were treated with DMARDs and 31 (58.5%) with biologicals.

Table 1

Characteristics	N= 53
Sex, female (%)	40 (75)
Age, years mean (SD)	58 (14)
Disease duration, months median (Min. – Max.)	57 (0 - 354)
CRP, mg/dL median (Min. –Max.)	0.40 (0.20 – 15.5)
ESR, mm median (Min. – Max.) (N= 50)	14 (1 – 80)
RF positivity (%)	31 (58)
Sharp- v/d Heijde score, median (Min. – Max.) (N= 50)	40 (3 – 186)
HAQ score, median (Min. – Max.)	0.88 (0 – 2.75)
DAS28, median (Min. – Max.) (N= 50)	2.89 (0.19 – 6.45)
Current therapy, number (%) NSAID DMARD Biological	 30 (57) 40 (75) 32 (60)

SD: Standard deviation, CRP: C-reactive protein, HAQ: Health Assessment Questionnaire, NSAID: Non-Steroidal Anti-Inflammatory Drugs, DMARD: Disease Modifying Anti-Rheumatic Drug, DAS28: Disease Activity Score of 28 joints.

Inter-observer reliability

For GSUS the ICC (A,1) ranged between 0.35 (95%CI:0.23–0.48) in position 3 and 0.60 (95%CI: 0.50–0.70) in the RC joint of position 1 and also 0.60 (95%: 0.50–0.70) in the IC joint of position 2. For PDUS the ICC (A,1) ranged between 0.36 (95%CI: 0.23– 0.49) in position 3 and 0.52 (95%CI: 0.41–0.63) in the RC joint of position 1 and 2. Details for each position are presented in Table 2.

The inter-observer reliability between the inexperienced ultrasonographer and one of the two experienced ultrasonographers with whom she had trained intensively ranged from 0.66 to 0.74 in GSUS and from 0.62 to 0.72 in PDUS. Whereas these values ranged from 0.39 to 0.40 in GSUS and from 0.20 to 0.47 in PDUS for the inexperienced ultrasonographer and the other experienced ultrasonographer with whom she had trained less.

Table 2

Interobserver reliability:

	Position 1	Position 2	Position 3
RC (GS)	0.60	0.52	0.35(UC)
IC (GS)	0.44	0.60	-
RC (PD)	0.52	0.52	0.36(UC)
IC (PD)	0.40	0.40	-

Right and left wrist together. Per examiner per joint.
(RC= radiocarpal joint, IC= intercarpal joint, GS= Greyscale, PD= Power Doppler, UC=ulnocarpal)

DISCUSSION

Wrist examination with US is made difficult due to the complex anatomy of the wrist and the amount of anatomical structures to evaluate. Therefore it is important to use distinct anatomical landmarks to describe probe positioning. Previous studies did not use these distinct landmarks, leaving the exact location where pathology was evaluated open for discussion. Using distinct anatomical landmarks improves reliable probe positioning providing consistent results between rheumatologists.

In this study with distinct bony landmarks, easy to locate on ultrasonographic examination, we showed that, when taking into account both radiocarpal and intercarpal joints, GSUS examination of the wrist for the presence of synovitis seems most reliable in position 2 (Lister's Tubercle to DIG III). For evaluating only the RC-joint, position 1 (Lister's Tubercle to DIG II) performed best. For PDUS, the inter-observer reliability was similar for position 1 and 2. The ulnocarpal orientation (position 3) showed the poorest reliability in both GSUS and PDUS. Although it is generally recognised that US underperforms in the wrist due to inter-examiner variation, only one study presented data on the reliability of only the wrist resulting in a kappa for PDUS of 0.69 (1). Most US reliability studies provide data for a combination of joints and not for the wrist separately (7, 10, 11). In our study PDUS reliability was fair. This may partly be explained by features of the ICC itself. Because point estimates of the ICC depend on the heterogeneity of the patients in the study it is necessary to have sufficient variation in the presence of pathology (17, 18). We selected RA patients with a heterogeneous clinical involvement of the wrist. This seemed not to be sufficient for the PD scores in our study. Many of our patients had a zero score on PD causing too little variation between patients, which is a reason for low ICCs (A,1). One way to cope with this in future studies is to select patients based on their PD status by an independent US examiner not involved in the reliability assessment.

To further improve inter-observer reliability in US wrist examination calibration between ultrasonographers is helpful. This was highlighted by the higher reliability between the two ultrasonographers who had trained most intensively together in this study. This same principle was described in an important MRI study where calibration exercises were used to improve inter-reader reliability of MRI scoring considerably (19). Another way to improve wrist evaluation is to scan healthy subjects getting familiar with the normal anatomical variation. This is especially true for the ulnocarpal joint in which the complex anatomy causes structures to appear pathologic due to their echogeneity. The moderate inter-observer reliability could be improved somewhat by using the multiplanar method, taking position 1 and 2 together. The ICC (A,1) would rise to 0.66 in both the RC and the IC-joint for GSUS and to 0.62 for PDUS in the RC-joint (data not presented).

In conclusion, US-examination of the wrist joints of RA patients with GSUS shows highest, moderate reliability using the anatomical landmarks Tubercle of Lister and digit III (position 2). The reliability for examination with PDUS is similar in both position 1 (Lister's Tubercle to digit II) and position 2 (Lister's Tubercle to digit III). Evaluation on the ulnar side (position 3) showed poor inter-observer reliability. This study suggests that position 2 should be used in clinical trials and in daily practice.

5

ACKNOWLEDGEMENTS

Note: Abbreviations are explained in the abbreviations list in the appendix.

REFERENCES

1. STRUNK J, HEINEMANN E, NEECK G, SCHMIDT KL, LANGE U: A new approach to studying angiogenesis in rheumatoid arthritis by means of power Doppler ultrasonography and measurement of serum vascular endothelial growth factor. *Rheumatology* (Oxford) 2004; 43: 1480-3.

2. COMBE B, LANDEWÉ R, LUKAS C *et al.*: EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2007; 66: 34-45.

3. SZKUDLAREK M, COURT-PAYEN M, 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.

4. WAKEFIELD RJ, D'AGOSTINO MA, IAGNOCCO A *et al.*: The OMERACT Ultrasound Group: status of current activities and research directions. *J Rheumatol* 2007; 34: 848-51.

5. CONAGHAN PG, MCQUEEN FM, PETERFY CG et al.: The evidence for magnetic resonance imaging as an outcome measure in proof-of-concept rheumatoid arthritis studies. *J Rheumatol* 2005; 32: 2465-9.

6. WAKEFIELD RJ, BALINT PV, SZKUDLAREK M *et al.*: Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol* 2005; 32: 2485-7.

7. KOSKI JM, SAARAKKALA S, HELLE M *et al.*: Assessing the intra- and inter-reader reliability of dynamic ultrasound images in power Doppler ultrasonography. *Ann Rheum Dis* 2006; 65: 1658-60.

8. HAAVARDSHOLM EA, OSTERGAARD M, EJBJERG BJ *et al.*: Reliability and sensitivity to change of the OMERACT rheumatoid arthritis magnetic resonance imaging score in a multireader, longitudinal setting. *Arthritis Rheum* 2005; 52: 3860-7.

9. TAOULI B, ZAIM S, PETERFY CG *et al.*: Rheumatoid arthritis of the hand and wrist: comparison of three imaging techniques. *AJR Am J Roentgenol* 2004; 182: 937-43.

10. NAREDO E, MOLLER I, MORAGUES C *et al.*: Interobserver reliability in musculoskeletal ultrasonography: results from a "Teach the Teachers" rheumatologist course. *Ann Rheum Dis* 2006; 65: 14-9.

11. NAREDO E, BONILLA G, GAMERO F, USON J, CARMONA L, LAFFON A: Assessment of inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with grey scale and power Doppler ultrasonography. *Ann Rheum Dis* 2005; 64: 375-81.

12. SIEGERT CE, VLEMING LJ, VANDENBROUCKE JP, CATS A: Measurement of disability in Dutch rheumatoid arthritis patients. *Clin Rheumatol* 1984; 3: 305-9.

13. PREVOO ML, VAN 'T HOF MA, KUPER HH, VAN LEEUWEN MA, VAN DE PUTTE LB, VAN RIEL PL: Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995; 38: 44-8.

14. STREINER DL NG: Health Measurement Scales. A Practical Guide To Their Develop-ment And Use. Oxford University Press; 2008.

15. MCGRAW K, WONG S: Forming inferences about some intraclass correlation coefficients. *Psycological Methods* 1996; 1: 30-46.

16. VON ELM E, ALTMAN DG, EGGER M *et al.*: The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008; 61: 344-9.
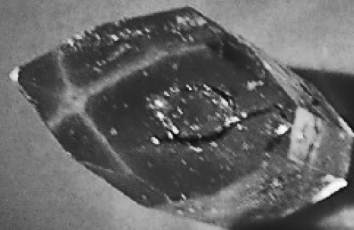
17.   TEN CATE DF, LUIME JJ, HAZES JM, JACOBS JW, LANDEWÉ R: Does the intraclass correlation coefficient always reliably express reliability? Comment on the article by CHEUNG *et al. Arthritis Care Res* (Hoboken) 2010; 62: 1357-1358; author reply 1358.

18.   LUKAS C, BRAUN J, VAN DER HEIJDE D *et al.*: Scoring inflammatory activity of the spine by magnetic resonance imaging in ankylosing spondylitis: a multireader experiment. *J Rheumatol* 2007; 34: 862-70.

19.   BIRD P, JOSHUA F, LASSERE M, SHNIER R, EDMONDS J: Training and calibration improve inter-reader reliability of joint damage assessment using magnetic resonance image scoring and computerized erosion volume measurement. *J Rheumatol* 2005; 32: 1452-8.

5

PART 2

Clinical domain
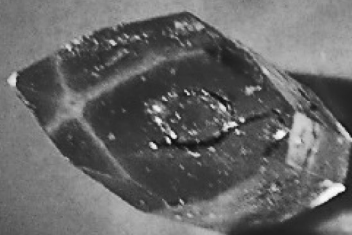
Monitoring of rheumatoid arthritis disease activity in individual patients: still a hurdle when implementing the treat-to-target principle in daily clinical practice.

Jacobs JW, Ten Cate DF, van Laar JM.

It is generally assumed that composite measures such as the DAS assessing 28 joints [DAS28; $0.7\ln(ESR) + 0.0142(VAS) + 0.555\sqrt{(28TJC)} + 0.284\sqrt{(28SJC)}$; VAS: visual analogue scale; TJC: tender joint count; SJC: swollen joint count] have been widely validated for use in individual RA patients. This assumption is false. The DAS28 has been validated for evaluations at group level, i.e. for measuring effects in clinical trials, but not for evaluations at individual patient level, i.e. for use in clinical practice. Is this only a theoretical, methodological issue or a relevant problem for daily practice?

The DAS28 has serious drawbacks, especially when used for applying the treat-to-target principle in an individual RA patient (1), which requires precise and valid measurement of disease activity. Joints of the ankles and feet are not included in the DAS28. At group level, this is not a major problem, but at the individual patient level, this may lead to misclassification regarding states of low disease activity or remission (2). One could argue, first, that in the absence of a generally accepted gold standard of remission, misclassification of remission more or less is arbitrary, and second, that according to the 2011 remission criteria, the absence of involved joints is not required (3). However, among RA patients in DAS28 remission, several have more than five swollen joints (SJs) and some more than 10 SJs, especially joints of ankles and feet (4). This cannot be easily ignored.

Issues with ankle and foot involvement aside, theoretically a patient can have 24 SJs of the 28 joints assessed for the DAS28 and still be in DAS28 remission. In contrast, a patient clinically in remission with four tender joints (TJs), zero SJs, a VAS global of 15 mm (range 0-100 mm) and an ESR of 15 mm/h is classed as having moderate disease activity according to the DAS28.

Furthermore, concomitant (secondary) FM or tender points in the non-FM range may result in a high DAS28 (5, 6). This is due to a positive association of tender points with the VAS global and TJ scores (6). This too might lead to misclassification, for instance, falsely rejecting remission in an RA patient who is clinically in remission. Concomitant FM is not rare; it is present in 12-17% of patients with RA (6).

Another issue arises when comparing individual patients' DAS28. Similar scores may reflect different grades of disease activity and stages of RA. A DAS28 of 3.9 (moderate disease activity) can be based on 8 TJs, 0 SJs, a VAS of 30 mm and an ESR of 15 mm/h, but also on 0 TJs, 12 SJs, a VAS of 30 mm and an ESR of 35 mm/h. There are also problems related to individual DAS28 components, in part caused by their relative weightings and conversions. Because in the DAS28 formula a TJ has a weight of 1.95 times that of a SJ and because square roots are applied, three TJs contribute more than seven SJs to the overall score.

ESR is not just a measure of acute phase reaction, it is also increased by conditions such as obesity (also associated with difficult clinical assessment of disease activity) and hyperlipidaemia. ESR is frequently slightly elevated in elderly patients. Would a
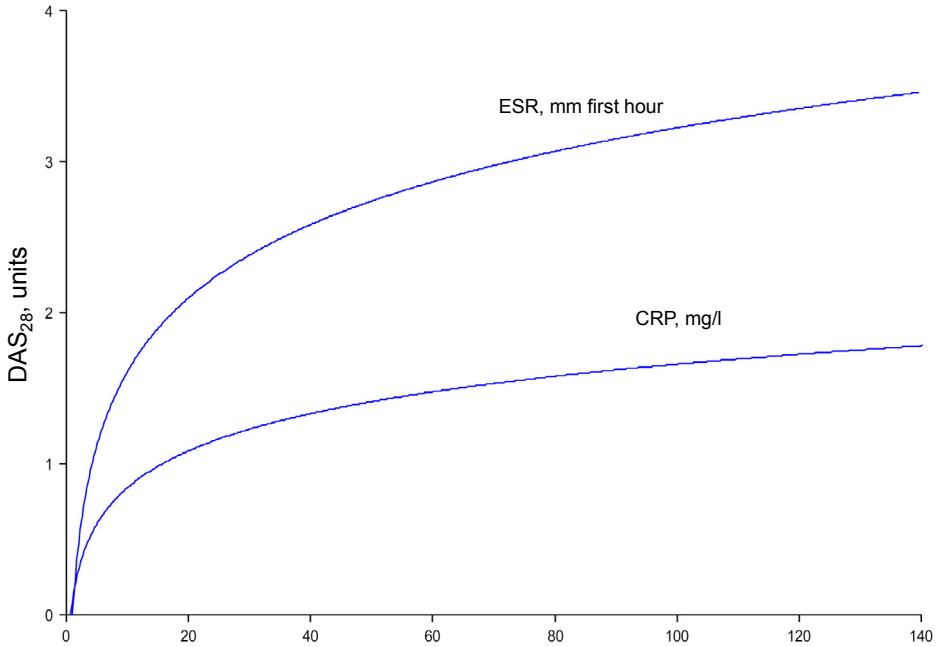
Figure 1. Plot of the contribution of ESR and CRP to the DAS28. Changes in the lower ranges of ESR and CRP influence the DAS28 most. DAS28: 28-joint DAS.

slight increase really matter? Well yes, because changes in the lower range of the ESR influence DAS28 most (Fig. 1), due to its log conversion. A non-specifically raised ESR might lead to a false conclusion that remission is absent. In comparison with ESR, CRP is less influenced by other conditions; it can be applied in the DAS28 instead of ESR, as $0.36\ln(CRP + 1)$, with CPR in milligrams per litre. However, CRP then has the same drawback as ESR, i.e. changes in its lower range influence DAS28 most (Fig. 1).

The clinical disease activity index and the simplified disease activity index have no weightings or arithmetic conversions of individual components and each includes a physician's global assessment. Therefore they are not subject to some of the flaws described above and also are much easier to calculate. However, they also assess only 28 joints (7).

Based on the issues described above, we decided not to use the DAS28 for individual patients' strategy steps in our treat-to-target computer-assisted management in early RA (CAMERA) trials (8, 9). The DAS28 and other indexes have too many flaws to apply them as sole determinants of RA activity for treat-to-target strategies or as criterion for access to biologic therapy. What is really needed is a composite measure of disease activity without the flaws mentioned above and validated for use in individual patients. Until that time we do not argue against the use of a composite

index in daily practice, or the scoring of joint counts for that matter. However, we do recommend that more than 28 joints should be assessed and that the individual components should be taken into account too. The scores should be interpreted in the light of particularities of the patient (e.g. age, BMI, co-morbidities and stage of the disease). For a patient with elevated ESR not related to disease activity, one could choose an index with CRP instead. To improve the specificity of assessing remission, one could add to the DAS28 criterion of remission the requirement of absence of any SJs, including ankles and feet. Alternatively, one might allow one SJ, according to the Boolean definition of the 2011 remission criteria (3). US may be applied to assist clinical decision making. Some US issues remain to be resolved. At the moment there is no consensus on clinically relevant cut-off scores and the selection of joints to scan. Nevertheless, power Doppler US has added value when diagnosing RA and when evaluating remission of RA (10). Another argument for applying US in clinical decision making in individual patients is that power Doppler signs of synovitis predict the progression of radiological joint damage and future flare in RA patients clinically in remission (10). In conclusion, monitoring of RA disease activity in daily clinical practice demands a personalized approach; clinical decision making based only on an aggregate value of a composite index such as DAS28 is insufficient.

Note: Abbreviations are explained in the abbreviations list in the appendix.

6

REFERENCES

1   Smolen JS, Aletaha D, Bijlsma JW *et al*. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 2010;69:631-7.

2   Bakker MF, Jacobs JW, Kruize AA *et al*. Misclassification of disease activity when assessing individual patients with early rheumatoid arthritis using disease activity indices that do not include joints of feet. Ann Rheum Dis 2012;71: 830-5.

3   Felson DT, Smolen JS, Wells G *et al*. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Arthritis Rheum 2011;63:573-86.

4   Landewe´ R, van der Heijde D, van der Linden S, Boers M. Twenty-eight-joint counts invalidate the DAS28 remission definition owing to the omission of the lower extremity joints: a comparison with the original DAS remission.Ann Rheum Dis 2006;65:637-41.

5   Ranzolin A, Brenol JC, Bredemeier M *et al*. Association of concomitant fibromyalgia with worse disease activity score in 28 joints, health assessment questionnaire, and short form 36 scores in patients with rheumatoid arthritis.Arthritis Rheum 2009;61:794-800.

6   Ton E, Bakker MF, Verstappen SM *et al*. Look beyond the disease activity score of 28 joints (DAS28): tender points influence the DAS28 in patients with rheumatoid arthritis. J Rheumatol 2012;39:22-7.

7   Gaujoux-Viala C, Mouterde G, Baillet A *et al*. Evaluating disease activity in rheumatoid arthritis: which composite index is best? A systematic literature analysis of studies comparing the psychometric properties of the DAS, DAS28, SDAI and CDAI. Joint Bone Spine 2012;79:149-55.

8   Verstappen SM, Jacobs JW, van der Veen MJ *et al*. Intensive treatment with methotrexate in early rheumatoid arthritis: aiming for remission. Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA, an open-label strategy trial). Ann Rheum Dis 2007;66: 1443-9.

9   Bakker MF, Jacobs JW, Welsing PM *et al*. Low-dose prednisone inclusion in a methotrexate-based, tight control strategy for early rheumatoid arthritis: a randomized trial. Ann Intern Med 2012;156: 329-39.

10 Ten Cate DF, Luime JJ, Swen N *et al*. Role of ultrasonography in diagnosing early rheumatoid arthritis and remission of rheumatoid arthritis—a systematic review of the literature. Arthritis Res Ther 2013; 15:R4.

CHAPTER 7

Role of ultrasonography in diagnosing early rheumatoid arthritis and remission of rheumatoid arthritis -- a systematic review of the literature.

Ten Cate DF, Luime JJ, Swen N, Gerards AH, De Jager MH, Basoski NM, Hazes JM, Haagsma CJ, Jacobs JW

ABSTRACT

Introduction: Ultrasonography (US) might have an added value to clinical examination in diagnosing early rheumatoid arthritis (RA) and assessing remission of RA. We aimed to clarify the added value of US in RA in these situations performing a systematic review.

Methods: A systematic literature search was performed for RA, US, diagnosis and remission. Methodological quality was assessed; the wide variability in the design of studies prohibited pooling of results.

Results: Six papers on the added value of US diagnosing early RA were found, in which at least bilateral metacarpophalangeal (MCP), wrists and metatarsophalangeal (MTP) joints were scanned. Compared to clinical examination, US was superior with regard to detecting synovitis and predicting progression to persistent arthritis or RA. Eleven papers on assessing remission were identified, in which at least the wrist and the MCP joints of the dominant hand were scanned. Often US detected inflammation in patients clinically in remission, irrespective of the remission criteria used. Power Doppler signs of synovitis predicted X-ray progression and future flare in patients clinically in remission.

Conclusions: US appears to have added value to clinical examination for diagnosing of RA when scanning at least MCP, wrist and MTP joints, and, when evaluating remission of RA, scanning at least wrist and MCP joints of the dominant hand. For both purposes primarily power Doppler US might be used since its results are less equivocal than those of greyscale US.

INTRODUCTION

The treatment of rheumatoid arthritis (RA) has improved dramatically over past decades with the early and intensive use of conventional disease modifying antirheumatic drug (DMARD) strategies (1) and the introduction of biological agents (2). Treatment strategies with dose and medication adjustments tailored to the individual patient (tight control) to achieve a predefined level of low disease activity, or preferentially, remission within a certain limited period of time (treat to target) (3) are nowadays used for early RA (4). It is widely accepted that early after the onset of RA, there is a period of time (window of opportunity) during which effective treatment can beneficially alter the outcome of the disease in the long term (5,6). This requires prompt referral and recognition of RA. Recently, new classification criteria (7) and new remission criteria (8) have been published. In the new classification criteria, it is suggested that imaging techniques such as ultrasonography (US) may be used for additional information in joints clinically suspected of arthritis (7). Regarding remission criteria, a considerable number of patients in clinical remission according to several clinical criteria shows signs of inflammation on US (9-11). These findings imply that US may have added value to clinical examination when diagnosing RA, or evaluating remission in RA. For this purpose we would have to make a selection of joints to evaluate by US, because assessment of all joints would be very time consuming. The aim of this systematic review is to clarify if US, when used for diagnosing RA and for evaluating remission in RA, would give additional information to clinical examination, to elucidate which minimal set of joints should be assessed by US for these two purposes, and by which modality, that is, power Doppler US (PDUS) and/ or greyscale US (GSUS).

7

MATERIAL AND METHODS

A systematic literature search was performed in PubMed, Embase and the Cochrane library for articles published up to October 4, 2011. A list of relevant keywords and synonyms for disease (RA, arthritis) and imaging (ultrasonography) was compiled. Keywords, including words of the title and abstract, and medical subject headings (Mesh) were combined using Boolean operators (AND, OR) (see Additional file 1). Included studies were those on adult humans, published in the English or Dutch language, either on diagnosing RA or evaluating US signs of synovitis in RA patients who were clinically in remission. We limited our study to the signs of US inflammation and did not assess structural joint damage. In the domain of diagnosis, assessing

structural change probably would not increase the additive value of US very much, given the already very sensitive current 2010 American College of Rheumatology (ACR) classification criteria. Second, there would have been the problem of how to apply the US finding of structural damage to the 2010 ACR classification criteria. In the domain of remission, structural changes are a sign of past inflammation only. Disagreements about study inclusion were resolved by discussion; results are based on full consensus. Excluded were reviews, editorials, case reports and letters to the editor. One reviewer (DTC) screened titles and abstracts. Relevant articles were obtained and their reference lists were screened to find additional studies. Data were extracted by one reviewer (DTC) on year of publication, study population, study design and duration, treatment, possible follow up, number and type of joints under investigation, statistical methods and US parameters.

We used an adaptation of the phases (levels) in diagnostic studies proposed by Sackett and Haynes to reflect the clinical relevance of research data (Additional file 2, box 1) (12). All results were summarized descriptively. Heterogeneity in study design and methods precluded pooling the results. Methodological quality of studies on diagnosing RA was assessed by the instrument Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) (13), with an extra question on sample size. For the studies evaluating remission a quality assessment tool was not available; we created a quality items list (Additional file 3). The quality assessments were performed to check for possible flaws in study design and analyses.

RESULTS

Diagnosing RA

In our systematic search we found six papers on the added value of US joint core sets in diagnosing RA (14-19). All studies can be considered phase three, according to Sackett and Haynes. Arthritis was evaluated using both GSUS and PDUS. For definitions of US signs of inflammation see Table 1. In four of these studies wrists and metacarpophalangeal (MCP) joints were assessed as the minimum (14-17); study-one also evaluated tendons (14), study-two also evaluated proximal interphalangeal (PIP) joints (15), and the third study extended the core set to metatarsophalangeal (MTP) joints and larger joints, such as the shoulders, knees and ankle (16). The fourth study also included the distal interphalangeal (DIP) and elbow joints (17). In the fifth study (18), painful joints and the adjacent joints of the same joint region (if applicable) and their contralateral joints were assessed, whereas the sixth study started scanning

painful joints only; during this study the protocol was changed to US of MCP joints, MTP joints and knees bilaterally (19).

At the joint level, among anti-cyclic citrullinated peptide (aCCP)-positive arthralgia patients (that is, those having no clinically swollen joints) US predicted progression to clinically detectable joint inflammation of the subset of joints showing a positive PDUS signal, after a median follow up of 26 months, with an odds ratio (OR) of 5.50 (95% CI 2.57, 11.9) (18). Another study showed that adding US parameters at baseline to clinical parameters increased the pretest probability of 6% to 94% posttest for the progression to inflammatory arthritis at 12 month follow up at patient level. This was evaluated among 30 rheumatoid factor (RF) and/or aCCP-negative individuals with inflammatory hand symptoms with or without clinical synovitis (14). Among individuals with possible RA (n = 58), 10% (three out of twenty-nine patients) were rightly classified as RA patients by US at baseline, using the clinical diagnosis of RA at 1.5 years as the reference standard (16). In 80 patients with early oligo-arthritis(< 12 months), about 1/3 of patients could be reclassified as having > 5 inflamed joints when US was added to the clinical examination, but 15 of all 185 joints (8%) with clinical synovitis were normal on US examination (19). This is why the study extended its scan protocol halfway through the study, from scanning only painful joints to scanning MCP and MTP joints and knees bilaterally. In one study in which the shoulders, elbows, wrists, MCP joints, PIP joints, DIP joints, knees, ankles and MTP joints were scanned, among 51 inflammatory arthritis patients, the subgroup with US symmetric polyarthritis was compared with the subgroup, who two years thereafter, met the ACR 1987 criteria for RA, yielding a kappa statistic of 0.61, which denotes a reasonable level of agreement (17).

In another study US at baseline reclassified 15% (n =22) of the patients with undifferentiated arthritis (n =149) to having RA, using the clinical diagnosis of RA (n = 62) with a follow up of 12 months as reference, while US in 11% (n = 17) was false positive at baseline. For this study, as the cut-off for US inflammation, a PDUS signal in more than three joints was used (15). This suggest a US sensitivity of 0.35 (95% CI 0.24, 0.48) and a specificity of 0.78 (95% CI 0.67, 0.86) for diagnosing RA at baseline.

Quality assessment of studies reviewed for diagnosing RA

Details of quality assessment of studies are shown in Table 1 and Additional file 3. Assessment of the methodological quality using the QUADAS-2 (13), extended with a question on sample size, showed relevant patients in all six studies, who were followed over time in five studies (15,16,18). Both the index test (US) and reference (diagnosis of RA) were clearly described and applied to all included patients, although

Table 1. Value of ultrasonography in diagnosing early (rheumatoid) arthritis

Author (year)[ref] Phase*	Study population	Number; type of joints or tendons	US modality, definition of synovitis (A-F)	Outcome, Risk estimates (95% CI)	Authors' conclusions	Limitations#
Filer (2011) [18] 3	58 IA 22 RA	38; MCP, PIP, MTP, wrists, elbows, shoulders, knees, ankles	GSUS & PDUS A	US reclassifies 3 of 29 IA patients to RA (ACR 2010)	US of MCP, wrists and MTP improves optimally clinical models to predict RA	Q4E
Freeston (2010) [17] 3	30 patients with inflammatory symptoms and negative RF and aCCP tests	12; MCP, wrists and flexor tendons of fingers	GSUS & PDUS B	US increases pretest probability for arthritis at one year from 6 to 94%	Inflammation at US predicts development of arthritis in patients with inflammatory symptoms and negative RF and aCCP tests	Q1D, Q2C
Ozgul (2009) [14] 3	51 IA	34; shoulders, elbows wrists, MCP, PIP, DIP, knees, ankles, MTP	GSUS C	Kappa = 0.61 for US diagnosis of RA versus clinical diagnosis	Average agreement between US and clinical examination when diagnosing RA	Q1A, Q1E, Q4E
Salaffi (2010) [19] 3	149 UA	18; wrists, MCP2-5, PIP2-5	GSUS & PDUS D	Sensitivity 0.35 (0.24, 0.48) and specificity 0.78 (0.67, 0.86) for US as baseline test, predicting RA at 1 year	Presence of PDUS signs predicts progression from UA to RA	Q1A, Q1C, Q1E, Q3B, Q4E
Stadt (2010) [16] 3	192 RF and/or aCCP-positive arthralgia patients, clinically without arthritis	Painful joint + adjacent and contralateral joints, mean 8 joints per patient	GSUS & PDUS E	JE: OR = 3.07 (1.05, 8.9); SH: OR = 5.5 (2.3, 12) PD: OR = 5.5 (2.6, 12), all for future arthritis	Inflammation at US predicts future arthritis in aCCP-positive patients clinically without arthritis	Q4E

Table 1. (continued)

Author (year)[ref] Phase*	Study population	Number; type of joints or tendons	US modality, definition of synovitis (A-F)	Outcome, Risk estimates (95% CI)	Authors' conclusions	Limitations#
Wakefield (2004) [15] 3	80, Oligoarthritis	First 40 patients: only symptomatic joints Last 40 patients: 22; MTP, MCP, Knees	GSUS F	1/3 of cases of oligoarthritis reclassified to polyarticular (>= 6 joints) arthritis	GSUS detects inflammation in more joints than clinical exam, but 8% of joints with clinical synovitis are normal at US	Q2C

A, according to Szkudlarek 2003 [30], 2006 [31] and Wakefield 2005 [32]. B, semiquantitative definitions based on binary definitions by Brown 2006 [11], Wakefield 2004 [15], Karim 2004 [33] and Newman 1996 [34]. C, (a) joint capsule elevated beyond normal range, or (b) synovial hypertrophy or effusion around the joints, or (c) erosions of the joint margins, or (d) change in the diameter of the tendon with or without peritendinous hypo-echogeneity supporting effusion. D, according to Wakefield 2005 [32] and Szkudlarek 2003 [30]. E, according to Szkudlarek 2003 [30]. F, abnormally hypo-echoic joint space, distinct from the intra-articular fat pad and non-compressible with the transducer. Ref, reference; US, ultrasonography; GSUS, greyscale ultrasonography; PDUS, power Doppler ultrasonography; aCCP, anti-cyclic citrullinated peptide; JE, joint effusion; IA, inflammatory arthritis; RA, rheumatoid arthritis; UA, undifferentiated arthritis; RF, rheumatoid factor; SH, synovial hypertrophy; OR, odds ratio. *Adaptation of phases according to Sackett and Haynes (Additional file 2, box 1): Phase 1. Do US results in patients with the condition differ from those without the condition? Phase 2. Are patients with certain US results more likely to have the condition? Phase 3. Do US results distinguish patients with and without the condition among those in whom it is clinically sensible to suspect the condition? Phase 4. Do patients undergoing US fare better in their ultimate health outcome than similar untested patients?. #See Additional file 3.

semiquantitative definitions of US inflammation used in one paper eventually seem to be based on a paper using a binary score and a paper describing synovitis of the knee (14). Drop out of patients was mentioned in one of five longitudinal studies (14).

Furthermore, in our interpretation small sample sizes and heterogeneity of studies diminished the strength of evidence of the value of US to improve early diagnosis of RA (14-17).

Evaluation of remission of RA

Our systematic search yielded 11 papers on the added value of US in the evaluation of remission in RA (9-11,20-27), using sets of joints ranging from six (10) to forty-four joints (26). One study scanned forty-four joints (26), two studies scanned forty-two (20,27), and the other studies assessed between six and sixteen joints (9-11,21-25). Within the 11 studies, the wrist and MCP joints of the dominant hand were always scanned. Arthritis was evaluated by GSUS and PDUS. The definition of remission varied and included physician-determined remission (11,20,21), a disease activity score (DAS) <1.6 (23,26), and complete absence of clinical and laboratory symptoms (9), while the time since remission varied from 2 months (9,20,22) to 3 months (26), or to more than 6 months (10,11,21,23-25). One study evaluated the time to remission in a treatment setting (27). Study characteristics, definitions of US signs of inflammation and outcomes are presented in Table 2.

In all 11 papers, there was a discrepancy between the number of clinically swollen joints and the higher number of joints with US signs of arthritis, indicating that joints that were not clinically inflamed showed US signs of arthritis. In five of these eleven papers it was explicitly mentioned that US evidence of synovitis was also found in joints that were not clinically inflamed (10,11,23,24,26). There seemed to be no clear association between the number of joints scanned per patient and the number of patients with at least one joint with US signs of synovitis. When looking at GSUS signs of synovitis, all 11 studies identified synovitis in 73 to 95% of patients in clinical remission; for PDUS signs of synovitis, the range was 8.7 to 62% (9-11,20-27).The predictive ability of US for clinical flares was evaluated in four studies (22,23,25,26). Three of these identified predictive value, where one did not (25). One study reported an OR of 3.6 (95% CI 1.4, 9.0) for the occurrence of flare in PDUS-positive patients when scanning the wrists, and the second and third MCP and PIP joints, all bilaterally (23). In another study, PDUS signs of synovitis were associated with an OR of 6.3 (95%CI 2.0, 20) (22) for the occurrence of flare within one year among patients in clinical remission (DAS44 < 2.4) when scanning the wrists, and the second, third and fifth MCP and MTP joints, all bilaterally. In a study assessing 44 joints among patients

in remission, PDUS signs of synovitis predicted flare with an OR of 13 (95%CI 1.6, 104) (26). The predictive value of GSUS was either not significant (22,25,26) or not presented (23).

Two of the eleven papers evaluated progression of radiological joint damage in patients in clinical remission. The presence of PDUS signs of inflammation increased the risk of joint damage with an OR of 1.4 (95%CI 1.1, 1.9) at the patient level in a study of nine patients with radiographic signs of progression (22). At the joint level, presence of PDUS signs of inflammation predicted progression with an OR of 12 (95%CI 3.3, 44) in a study of 10 patients (21). GSUS scores were significantly higher in the group that progressed vs. the group that did not progress (mean 4.8 ± SD 2.3 vs. 3.2 ± 2.6) (22), or they predicted radiographic progression with an OR of 1.92 (95% CI 0.49, 7.24) (21).

The impact of using different remission criteria was reported in five studies. Two of these studies presented discrepancies between the prevalence of inflammation detected by US if applying different remission criteria. The first of these two studies showed that among patients in remission according to the simplified disease activity index when using a cutoff point of less than 3.3 (SDAI score < 3.3), the number of joints with PDUS signs of synovitis was smaller, and the PDUS grade of synovitis was lower when compared to those of patients in remission when using a cutoff point of less than 5.5 on the SDAI (SDAI score < 5.5), the DAS28 (cutoff score < 2.4 or < 2.6) or slightly modified ACR 1981 remission criteria (that is, excluding the fatigue criterion). In this study 42 joints were scanned (20). The second of these two studies showed that using the ACR 1981 remission criteria, the number of patients with US inflammation was smaller compared to applying the criterion of a DAS < 1.6 (23). In this study 12 joints in the hands and wrists were scanned. The three other studies reporting on the impact of using different remission criteria showed that 60 to 80% of their patients had GSUS signs of inflammation independently of the specific criteria used (9-11). Regarding PDUS signs of inflammation, two of these three studies showed that these signs were present in about 50% of the patients in clinical remission when scanning the MCP joints and the wrist of the dominant hand (10,11). The third study, in which a greater number of joints was scanned, showed similar results: about 60% of the patients in clinical remission showed, irrespective of the clinical remission criteria used, PDUS signs of inflammation when scanning the wrists (ulnar and radial styloid regions) and the first to fifth MCP joints, all bilaterally (9).

The influence of disease duration was studied in one paper, assessing 12 joints in the hands and wrists. Among patients who were in clinical remission, 44% of those with early RA had no US signs of synovitis (defined as absence of GSUS and PDUS signs of inflammation) vs. 17% of those with longstanding RA. When defining synovitis as presence of GSUS synovitis and absence of PDUS synovitis, 15% of those with early

Table 2: Value of ultrasonography in remission of rheumatoid arthritis

Author [year][ref] Phase*	Study population, duration of remission or LDA (definition)	Number; type of joints	US modality, definition of synovitis (A-F)	Outcome¶ Risk estimates (95% CI)	Authors' conclusions	Limitations#
Balsa (2010) [20] 2	97 RA in remission >2 months (physician's judgment); 88 controls	42; PIP, MCP, wrists, elbows, shoulders, knees, ankles, MT, MTP	GSUS & PDUS A	Of RA in remission, 92% GSUS-positive and 42% PDUS-positive; 88% of non-arthritic controls GSUS-positive	Of sets ARA 1981, DAS28 and SDAI, the SDAI is closest to the absence of inflammation on US	j,o
Brown (2006) [11] 2	107 RA in remission > 6 months (physician's judgment)	8; MCP2-5, wrist (RC, IC, DRU, UC) of dominant hand	GSUS & PDUS B	73% GSUS-positive 43% PDUS-positive	In many patients in remission, inflammation is found on US	o,x
Brown (2008) [21] 2	102 RA in remission > 6 months (physician's judgment)	8; MCP2-5, wrist (RC, IC, DRU, UC) of dominant hand	GSUS & PDUS B	PDUS predicts radiographic progression at one year: OR = 12 (3.3, 44)	Presence of PDUS predicts radiographic progression in joints of patients in remission	o,v,x
Foltz (2011) [22] 2	85 RA in remission or with LDA > 2 months (DAS < 2.4)	14; wrists, MCP2,3,5, MTP2,3,5	GSUS & PDUS A	PDUS predicts flare: OR = 6.3 (2.0, 20); PDUS predicts radiographic progression: OR = 1.4 (1.1, 1.9)	Inflammation on PDUS predicts radiographic progression and flare of patients in remission or with LDA	o,q,v
Ozgocmen (2008) [9] 1	52 RA of whom a subgroup in remission > 2 months (Preliminary ARA 1981 + DAS < 3.2)	14; MCP1-5, wrists (USTL, RSTL)	GSUS & PDUS C	62% of 32 patients (in ARA 1981 remission) PDUS-positive 58% of 31 patients (in DAS28 remission) PDUS-positive	PDUS inflammation is found for 2 different remission criteria sets, in a similar number of patients	i,j,k,l,o,r,t,

Table 2. (continued)

Author (year)[ref] Phase*	Study population, duration of remission or LDA (definition)	Number; type of joints	US modality, definition of synovitis (A-F)	Outcome¶ Risk estimates (95% CI)	Authors' conclusions	Limitations#
Peluso (2011) [23] 2	48 early (7 months) RA in remission > 6 months (DAS < 1.6) 46 LSRA (118 mo.) in remission > 6 months (DAS < 1.6)	12; MCP2-3, PIP2-3, wrists (RC, UC)	GSUS & PDUS Not given	56% early RA GSUS-positive; 83% LSRA GSUS-positive; PDUS predicts flare: OR =3.6 (1.4, 9.0)	PDUS shows predictive value for flare In many patients in remission, inflammation is found on US	g,i,l,n,o,p,r,t,u,v
Saleem (2009) [24] 2	100 RA in remission > 6 months (DAS < 2.6)	16; wrists (RC, UC,IC), MCP1-5	GSUS & PDUS D	88% GSUS-positive PDUS-positive 58%	In >55 % of patients in remission inflammation is found on PDUS	i,o,r,x
Saleem (2010) [25] 2	47 RA in remission > 6 months (DAS28 < 2.6) of whom a subgroup of 20 patients is assessed also by US	18; wrists, PIP, MCP	GSUS & PDUS F	In group that flared after 2 years, at baseline 88% GSUS-positive and 44% PDUS-positive. In group that did not flare after 2 years, at baseline 82% GSUS-positive and 45% PDUS- positive	In many patients in remission inflammation is found on PDUS, but does not predict flare	o,r,u,v,wx
Saleem (2011) [10] 2	128 RA in remission > 6 months (ACR 2011)	6; wrist, MCP2-5 of the dominant hand	GSUS & PDUS D	Around 50% PDUS-positive when using ACR2011, DAS28, SDAI remission sets	For 3 different remission criteria sets a similar number of patients is PDUS-positive	c,h,i,j,k,l,o,p,r,t,u,v,x

7

Table 2. (continued)

Author (year)[ref] Phase*	Study population, duration of remission or LDA (definition)	Number; type of joints	US modality, definition of synovitis (A-F)	Outcome[¶] Risk estimates (95% CI)	Authors' conclusions	Limitations[#]
Scire (2009) [26] 2	106 early RA (<12 months), of whom 43 in remission > 3 months (DAS < 1.6)	44; shoulders, elbows, wrists, MCP, PIP, knees, ankles, MTP	GSUS & PDUS E	95% GSUS-positive, 41% PDUS-positive PDUS positivity predicts flare: OR = 13 (1.6, 104)	In many patients in remission, inflammation found on PDUS, predictive of flare	g,i,j,k,l,m,p,r,t,u,v
Wakefield (2007) [27] 3	10 early (<12 months) RA in remission (DAS28 < 2.6)	42; shoulders, elbows, wrists, MCP, PIP, knees, tibiotalar, MT, MTP	GSUS & PDUS Not given	50% GSUS-positive and 15% PDUS-positive in clinically normal joints	In many RA-patients in remission, PDUS inflammation is found	i,n,t,x

A, according to Wakefield 2005 [32]. B, semi quantitative definitions based on binary definitions by Brown 2006 [11], Wakefield 2004 [15], Karim 2004 [33] and Newman1996 [34]. C, GSUS: Hypoechogenity; PDUS: grade 0, no flow signal in the synovium; grade 1, separate dot signals or short linear signals; grade 2, clearly discernible vascularity with either many small vessels or several long vessels with or without visible branching, though involving less than half the area of the synovium; grade 3, vessels involving more than half the area of the synovium. D, semiquantitative definitions based on binary definitions by Brown 2006 [11], Wakefield 2004 [15], Karim 2004 [33]. No PDUS definitions. E, Wakefield 2005 [32] and Newman 1996 [34]. F, according to Outcome Measures in Rheumatology (OMERACT) definitions, but not clear from footnotes which definitions have been used. US, ultrasonography; GSUS, greyscale ultrasonography; PDUS, power Doppler ultrasonography; aCCP, anti-cyclic citrullinated peptide; JE, joint effusion; SH, synovial hypertrophy; IA, inflammatory arthritis; LDA, Low disease activity; LSRA, long standing RA; RA, rheumatoid arthritis; UA, undifferentiated arthritis; RF, rheumatoid factor; OR, odds ratio; DRU, distal radio-ulnar; IC, intercarpal; MT, midtarsal; RC, radiocarpal; RSTL, radial styloid; USTL, ulnar styloid; UC, ulnocarpal; SDAI, simplified disease activity index. *Adaptation of phases according to Sackett and Haynes (Additional file 2, box 1): Phase 1. Do US results in patients with the condition differ from those without the condition? Phase 2. Are patients with certain US results more likely to have the condition? Phase 3. Do US results distinguish patients with and without the condition among those in whom it is clinically sensible to suspect the condition? Phase 4. Do patients undergoing US fare better in their ultimate health outcome than similar untested patients? #See Additional file 3. [¶]GSUS-positive (%) and PDUS-positive (%) signify the percentages of individuals with at least one inflamed joint applying the respective US modality.

RA had US signs of synovitis, as opposed to 52% of the patients with longstanding RA. When defining synovitis as presence of both GSUS and PDUS signs of inflammation, 42% of the patients with early RA showed signs of US inflammation, as opposed to 30% of those with longstanding RA (23).

Quality assessment of studies reviewed for evaluation of remission of RA

Details of quality assessment are shown in Table 2 and Additional file 3. Quality assessment was performed for all studies but longitudinal studies are the most relevant to evaluate the remission of RA. Two of these studies evaluated the added value of US for prediction of radiographic progression among patients in remission, and four studies evaluated the value for prediction of flare. We found wide CIs and point estimates differing from study to study, probably due to small sample sizes, slightly different definitions for remission and flare and analyses performed at joint level and at patient level. Definitions of US inflammation are not clearly described in three of the eleven papers (23,25,27). In four other longitudinal studies the semiquantitative definitions of US inflammation used for inflammation in MCP joints and wrists seem to be based on a paper using a binary score and a paper describing synovitis of the knee (10,11,21,24).

7

DISCUSSION

The results of our systematic search indicate that when diagnosing RA a greater number of inflamed joints per patient was detected by US compared to clinical examination in populations ranging from aCCP/RF-positive patients with arthralgia, to patients with clinically observed arthritis. The presence of US signs of inflammation seems to increase the risk of progression to persistent arthritis or RA, implying clinical relevance. Regarding assessment of remission, our review shows that in many patients with low disease activity or in clinical remission, US signs of inflammation were detected, even in those who met stringent clinical remission criteria. These findings are relevant, because the results of these studies suggest that PDUS signs of synovitis predict progression of radiographic joint damage and flare. We limited our study to signs of US inflammation and did not assess structural joint damage. The reason for this is that in the publication of the new classification criteria, in which erosions are not included, it was suggested that US may be used to confirm clinical findings, i.e. swelling of the joint. Erosions typical of RA would imply the classification

of RA in patients who met the new classification criteria in the past. However, the new classification criteria are very sensitive: the diagnosis of RA can be made on the basis of one swollen joint, so one could argue that for these new criteria, the finding of erosions would not add much to the sensitivity in early RA, in contrast to the situation with the 1987 criteria. Second, there would have been the problem of how to apply the finding of structural damage, assessed by US, to the 2010 ACR classification criteria, that is, what would be the contribution of structural damage assessed by US to the diagnosis, applying the 2010 criteria? In the domain of remission, structural damage reflects inflammation in the past, not the current inflammatory state.

An important question is, which joints should be scanned? Scanning only the joints that are painful or clinically show arthritis does not seem to be a valid strategy, and scanning all joints is not feasible in daily practice. Based on the spectrum of joints most frequently involved in early RA and the results of this review, a recommendation when scanning for diagnosis of early RA could be to scan at the minimum the wrists, MCP and MTP bilaterally using PDUS; PIP joints could be included based on the results of one study. Also in the domain of remission of RA it is important to identify which joints to scan. Although more signs of arthritis are found when scanning a larger number of joints, a clear relation between the number of joints scanned and the number of patients clinically in remission with US signs of synovitis seems lacking. Therefore, it might be sufficient to scan a limited set of joints for this purpose. In eleven studies the wrist and MCP joints of the dominant hand had been scanned as the minimum.

Based on the results of this review it seems that it is not necessary to scan large joints when diagnosing RA or evaluating the remission of RA. In general, the more joints that are scanned, the higher the chance of finding US signs of arthritis in a patient. An earlier diagnosis leads to earlier initiation of adequate therapy, more often within the window of opportunity. This not only improves the prognosis in the short term, for example, by inducing remission at an earlier stage and more frequently, but possibly it also favorably alters the longterm course of the disease.

Another important question is which modality to use, PDUS or GSUS? Our systematic search indicates that PDUS in particular may have an added value in the diagnosis of early RA and evaluation of the remission of RA: the predictive value of PDUS was higher than that of GSUS. This is in line with the findings that GSUS signs of inflammation also occur in non-arthritic individuals (20). In a study in an osteoporosis outpatient clinic, GSUS signs of synovitis were detected in up to 88% of 16 individuals who were without clinical symptoms or signs of joint disease (controls), based on scanning 42 joints with a cutoff of at least one joint with a score of 1 according to the OMERACT criteria for synovial hypertrophy (20). Of all 672 joints scanned, 76 joints showed GSUS signs of synovitis, 64 of them with grade 1, 12 with grade 2, and none with grade 3 signs.

In another study, in which a total of 84 joints was scanned among nine healthy individuals, 23 joints showed GSUS grade 1 signs of inflammation and only one joint was scored grade 2; no joints had a grade 3 score (18). It seems that for the purpose of discriminating arthritis patients from non-arthritic patients, the use of the GSUS grade 1 score is debatable. Also in RA patients it is not clear what the significance of GSUS grade 1 is. One study states that in longstanding RA, GSUS might depict chronically thickened tissue without inflammation (23). At the patient level, a cumulative GSUS score for discriminating arthritis patients from non-arthritic patients has yet to be determined. A cutoff of 8 when scanning 22 joints has been proposed (28).

Although the predictive value of PDUS is higher than that of GSUS to predict early RA, flare of RA and radiographic progression, PDUS has limitations as well. It is a technique that is particularly operator-, machine- and setting-dependent (29). It is important to avoid pressure on the transducer, and motion artifacts, and to use the correct US settings, for example, wall filter and pulse repetition frequency should be low when assessing joints.

Although the study results in our systematic review generally were not conflicting for either diagnosis or remission, some considerations need to be made. For instance, the number of diagnostic studies is currently limited, and only one study has focused on the ACR/European League Against Rheumatism (EULAR) 2010 criteria. Furthermore, regarding the quality of studies, the longitudinal studies looking at events (flares or radiographic progression) are small, causing a wide variation in the US risk estimates. Also, the variables that have been shown to be predictors of the diagnosis of RA or of remission, such as radiographic joint data and aCCP test results, have not all been taken into account. This might have inflated the added value of US. In addition, clear definitions for US signs of inflammation were not always given.

Something else to consider is that some of the papers reviewed are from the same group (10,11,19,21,24,25,27). Data presented in these papers might not be independent of each other, with correlated results being biased in one or the other direction. However, the results from the studies in our review are based on different patient populations. Also, we did not find signs that this group may be evidently pro or contra US, such that it would affect their scientific integrity, especially since one of the papers from this group shows a lower predictive value of US compared to those in other reviewed papers. Large prospective longitudinal studies are necessary to evaluate the additional value of US in diagnosing RA, scanning joints and evaluating the predictive validity of other signs such as US-detected tenosynovitis.

7

CONCLUSIONS

In conclusion, although further research is needed, PDUS has additional value to clinical examination both in improving early diagnosis of RA and establishing true RA remission. GSUS seems less specific. In the diagnostic process, studies suggest that as a minimum the wrist, MCP and MTP joints should be scanned bilaterally, while for remission, studies suggest that as a minimum the wrist and MCP joints of the dominant hand should be scanned.

Note: Abbreviations are explained in the abbreviations list in the appendix.

REFERENCES

1. Lard LR, Visser H, Speyer I, vander Horst-Bruinsma IE, Zwinderman AH, Breedveld FC, Hazes JM: Early versus delayed treatment in patients with recent-onset rheumatoid arthritis: comparison of two cohorts who received different treatment strategies. *Am J Med* 2001, 111:446-451.

2. Bathon JM, Martin RW, Fleischmann RM, Tesser JR, Schiff MH, Keystone EC, Genovese MC, Wasko MC, Moreland LW, Weaver AL, Markenson J, Finck BK: A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. *N Engl J Med* 2000, 343:1586-1593.

3. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, Kincaid W, Porter D: Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004, 364:263-269.

4. Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, Combe B, Cutolo M, de Wit M, Dougados M, Emery P, Gibofsky A, GomezReino JJ, Haraoui B, Kalden J, Keystone EC, Kvien TK, McInnes I, MartinMola E, Montecucco C, Schoels M, van der Heijde D, Committee TTE: Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010, 69:631-637.

5. Quinn MA, Emery P: Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. *Clin Exp Rheumatol* 2003, 21(Suppl 31):S154-157.

6. Mottonen T, Hannonen P, Korpela M, Nissila M, Kautiainen H, Ilonen J, Laasonen L, Kaipiainen-Seppanen O, Franzen P, Helve T, Koski J, Gripenberg-Gahmberg M, Myllykangas-Luosujarvi R, Leirisalo-Repo M, therapy FI-RTGFRAC: Delay to institution of therapy and induction of remission using single-drug or combination-disease-modifying antirheumatic drug therapy in early rheumatoid arthritis. *Arthritis Rheum* 2002, 46:894-898.

7. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, Birnbaum NS, Burmester GR, Bykerk VP, Cohen MD, Combe B, Costenbader KH, Dougados M, Emery P, Ferraccioli G, Hazes JM, Hobbs K, Huizinga TW, Kavanaugh A, Kay J, Kvien TK, Laing T, Mease P, Menard HA, Moreland LW, Naden RL, Pincus T, Smolen JS, Stanislawska-Biernat E, Symmons D, *et al*: 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010, 69:1580-1588.

8. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, Aletaha D, Allaart CF, Bathon J, Bombardieri S, Brooks P, Brown A, Matucci-Cerinic M, Choi H, Combe B, de Wit M, Dougados M, Emery P, Furst D, Gomez-Reino J, Hawker G, Keystone E, Khanna D, Kirwan J, Kvien TK, Landewe R, Listing J, Michaud K, Martin-Mola E, Montie P, *et al*: American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis Rheum* 2011, 63:573-586.

9. Ozgocmen S, Ozdemir H, Kiris A, Bozgeyik Z, Ardicoglu O: Clinical evaluation and power Doppler sonography in rheumatoid arthritis: evidence for ongoing synovial inflammation in clinical remission. *South Med J* 2008, 101:240-245.

10. Saleem B, Brown AK, Keen H, Nizam S, Freeston J, Wakefield R, Karim Z, Quinn M, Hensor E, Conaghan PG, Emery P: Should imaging be a component of rheumatoid arthritis remission criteria? A comparison between traditional and modified composite remission scores and imaging assessments. *Ann Rheum Dis* 2011, 70:792-798.

11. Brown AK, Quinn MA, Karim Z, Conaghan PG, Peterfy CG, Hensor E, Wakefield RJ, O'Connor PJ, Emery P: Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying

antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. *Arthritis Rheum* 2006, 54:3761-3773.

12.   Sackett DL, Haynes RB: The architecture of diagnostic research. *BMJ* 2002, 324:539-541.

13.   Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM, Group Q: QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011, 155:529-536.

14.   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, 69:417-419.

15.   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* 28:686-694.

16.   Filer A, De Pablo P, Allen G, Nightingale P, Jordan A, Jobanputra P, Bowman S, Buckley CD, Raza K: Utility of ultrasound joint counts in the prediction of rheumatoid arthritis in patients with very early synovitis. *Ann Rheum Dis* 70:500-507.

17.   Ozgul A, Yasar E, Arslan N, Balaban B, Taskaynatan MA, Tezel K, Baklaci K, Ozguven MA, Kalyon TA: The comparison of ultrasonographic and scintigraphic findings of early arthritis in revealing rheumatoid arthritis according to criteria of American College of Rheumatology. *Rheumatol Int* 2009, 29:765-768.

18.   van de Stadt LA, Bos WH, Meursinge Reynders M, Wieringa H, Turkstra F, van der Laken CJ, van Schaardenburg D: The value of ultrasonography in predicting arthritis in auto-antibody positive arthralgia patients: a prospective cohort study. *Arthritis Res Ther* 2010, 12:R98.

19.   Wakefield RJ, Green MJ, Marzo-Ortega H, Conaghan PG, Gibbon WW, McGonagle D, Proudman S, Emery P: Should oligoarthritis be reclassified? Ultrasound reveals a high prevalence of subclinical disease. *Ann Rheum Dis* 2004, 63:382-385.

20.   Balsa A, de Miguel E, Castillo C, Peiteado D, Martin-Mola E: Superiority of SDAI over DAS-28 in assessment of remission in rheumatoid arthritis patients using power Doppler ultrasonography as a gold standard. *Rheumatology (Oxford)* 2010, 49:683-690.

21.   Brown AK, Conaghan PG, Karim Z, Quinn MA, Ikeda K, Peterfy CG, Hensor E, Wakefield RJ, O'Connor PJ, Emery P: An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. *Arthritis Rheum* 2008, 58:2958-2967.

22.   Foltz V, Gandjbakhch F, Etchepare F, Rosenberg C, Tanguy ML, Rozenberg S, Bourgeois P, Fautrel B: Power doppler but not low-field MRI predict relapse and radiographic disease progression in rheumatoid arthritis patients with low disease activity. *Arthritis Rheum* 2012, 64:67-76.

23.   Peluso G, Michelutti A, Bosello S, Gremese E, Tolusso B, Ferraccioli G: Clinical and ultrasonographic remission determines different chances of relapse in early and long standing rheumatoid arthritis. *Ann Rheum Dis* 2011, 70:172-175.

24.   Saleem B, Brown AK, Keen H, Nizam S, Freeston J, Karim Z, Quinn M, Wakefield R, Hensor E, Conaghan PG, Emery P: Disease remission state in patients treated with the combination of tumor necrosis factor blockade and methotrexate or with disease-modifying antirheumatic drugs: A clinical and imaging comparative study. *Arthritis Rheum* 2009, 60:1915-1922.

25.   Saleem B, Keen H, Goeb V, Parmar R, Nizam S, Hensor EM, Churchman SM, Quinn M, Wakefield R, Conaghan PG, Ponchel F, Emery P: Patients with RA in remission on TNF blockers: when and in whom can TNF blocker therapy be stopped? *Ann Rheum Dis* 2010, 69:1636-1642.

26. 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, 48:1092-1097.

27. Wakefield RJ, Freeston JE, Hensor EM, Bryer D, Quinn MA, Emery P: Delay in imaging versus clinical response: a rationale for prolonged treatment with anti-tumor necrosis factor medication in early rheumatoid arthritis. *Arthritis Rheum* 2007, 57:1564-1567.

28. Zufferey PMB, Brulhart L, Tamborini G, Scherer A, Ziswiler H: Persistence of ultrasound synovitis in the patients fullfilling the DAS and/or the new ACR/EULAR RA remission definitions: Results of the SONAR score applied to the patients of the SCQM cohort. *Ann Rheum Dis* 2012, 71(Suppl3):149.

29. Torp-Pedersen ST, Terslev L: Settings and artefacts relevant in colour/ power Doppler ultrasound in rheumatology. *Ann Rheum Dis* 2008, 67:143-149.

30. Szkudlarek M, Court-Payen M, Jacobsen S, Klarlund M, Thomsen HS, Ostergaard M: Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. *Arthritis Rheum* 2003, 48:955-962.

31. Szkudlarek M, Klarlund M, Narvestad E, Court-Payen M, Strandberg C, Jensen KE, Thomsen HS, Ostergaard M: Ultrasonography of the metacarpophalangeal and proximal interphalangeal joints in rheumatoid arthritis: a comparison with magnetic resonance imaging, conventional radiography and clinical examination. *Arthritis Res Ther* 2006, 8:R52.

32. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostino MA, Sanchez EN, Iagnocco A, Schmidt WA, Bruyn GA, Kane D, O'Connor PJ, Manger B, Joshua F, Koski J, Grassi W, Lassere MN, Swen N, Kainberger F, Klauser A, Ostergaard M, Brown AK, Machold KP, Conaghan PG, Group OSI: Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol* 2005, 32:2485-2487.

33. Karim Z, Wakefield RJ, Quinn M, Conaghan PG, Brown AK, Veale DJ, O'Connor P, Reece R, Emery P: Validation and reproducibility of ultrasonography in the detection of synovitis in the knee: a comparison with arthroscopy and clinical examination. *Arthritis Rheum* 2004, 50:387-394.

34. Newman JS, Laing TJ, McCarthy CJ, Adler RS: Power Doppler sonography of synovitis: assessment of therapeutic response—preliminary observations. *Radiology* 1996, 198:582-584.

7

Additional value of baseline ultrasonography in early RA to predict failure to reach DAS 28 remission after 1 year of tight-control treatment

Ten Cate DF, Jacobs JWG, Swen WAA, Hazes JMW, De Jager MH, Basoski NM, Haagsma CJ, Luime JJ*, Gerards AH*

*Both authors contributed equally
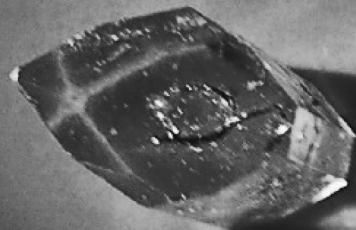
Assessment of disease activity in patients with rheumatoid arthritis using optical spectral transmission measurements, a non-invasive operator-independent imaging technique

Onna van M, Ten Cate DF, Tsoi KL, Meier AJL, Jacobs JWG, Westgeest AAA, Meijer PBL, Beek van MC, Rensen WHJ, Bijlsma JWJ

ABSTRACT

Objectives: In rheumatoid arthritis (RA), treat-to-target strategies require instruments for valid detection of joint inflammation. Therefore, imaging modalities are increasingly used in clinical practice. Optical spectral transmission (OST) measurements are fast, applicable during outpatient visits, and, in contrast to ultrasonography and the 28-joint Disease Activity Score (DAS28), operator independent and quantitative. We tested whether OST could measure disease activity validly in patients with RA.

Methods: In 59 patients with RA and 10 subjects with arthralgia, joint counts, DAS28, ultrasound (US) and OST examinations were performed. Additionally, MRI was performed in patients with DAS28<2.6.

Results: At the joint level, OST and US performed similarly regarding detection of inflammation in PIP-joints (area under the receiver-operating curve (AUC) of 0.79, p<0.0001) and MCP-joints (AUC 0.78, p<0.0001). Performance was less similar in wrists (AUC 0.62, p=0.018). On the patient level, OST correlated with clinical examination (DAS28 r=0.42, p=0.001), and US scores (r=0.64, p<0.0001). Further, also in patient with subclinical and low disease activity, there was a correlation between OST and MRI synovitis score (RAMRIS synovitis), r=0.52, p=0.005).

Conclusions: Operator-independent OST performed well in the detection of joint inflammation in patients with RA.

INTRODUCTION

The prognosis of rheumatoid arthritis (RA) has greatly improved with treatment targeting at remission or low disease activity (1). This requires instruments for valid detection of disease activity. Generally, the 28-joint Disease Activity Score (DAS28) is used for this. However, DAS28 has not been validated for use in individual patients, has considerable test-retest variability (2,3) and is also influenced by several factors unrelated to joint inflammation (4,5). Imaging techniques such as ultrasonography (US) and magnetic resonance imaging (MRI) are more sensitive for the detection of synovitis compared with physical assessment,(6) but they are not applied as standard assessments yet (7). US requires a skilled operator, is usually scored semiquantitatively and investigating multiple joints may be time-consuming. MRI with gadolinium enhancement is invasive, costly, time consuming and carries a risk of adverse effects and its use is therefore (except for clinical trials) usually restricted to the examination of a few joints. There is an unmet need for objective, fast measurement of disease activity at low cost, applicable during outpatient visits.

Optical spectral transmission (OST) measurement may be a candidate tool for non-invasive, fast and operator-independent measurement of disease activity. In OST, transmission of light of specific wavelengths is quantitatively measured. The same technique is used in pulse-oximetry (8). In the presence of synovitis, the transmission of light decreases (9,10), which is due to vascular changes (11). There is one commercially available optical device which measures reflection of light on illumination, before and after intravenously administrated indocyanine green (fluorescence optical imaging) (12,13). OST measures the transmission of light before and after temporarily impeding the venous return of blood from the forearms. Using this approach, OST through a single joint (proximal interphalangeal (PIP)) has shown to correlate with the detection of synovitis by physical examination (14). A multi-joint OST device (Full Hand Prototype (FHP)) has been developed, which was used in our current study. We tested whether the FHP could measure disease activity accurately in patients with RA.

9

METHODS

Subjects

Sixty patients with RA and ten controls with non-inflammatory arthralgia were recruited at the outpatient clinics of the departments of rheumatology at the University

Medical Center Utrecht (UMCU) and Regional Rheumatology Center Eindhoven. The aim was to include 20 patients for each category of disease activity, DAS28 < 2.6, 2.6 ≤ DAS28 ≥ 5.1, DAS28 > 5.1.

Preliminary stratification during recruitment was based on the DAS28 at the time of the screening visit at the outpatient clinic. Final stratification was based on the DAS28 measured at the day of examination, which resulted in 20 patients with DAS28 < 2.6, 27 with 2.6 ≤ DAS28 ≤ 5.1 and 13 with DAS28 > 5.1. Time between recruitment and enrolment was less than three weeks in all cases. Exclusion criteria were age less than 18 or over 90 years, significant deformities of the hand, wheelchair dependency, surgery of the wrist or hand in the preceding three months, light hypersensitivity, pregnancy and/or breastfeeding. In patients with DAS28 < 2.6 at recruitment, MRI was performed. In this subgroup only, additional exclusion criteria were applied: allergy to gadolinium, presence of MRI incompatible metal objects in any part of the body, renal insufficiency (defined as MDRD <30ml/min/1.73m2) and/or claustrophobia. All measurements were performed within a window of four hours at the University Medical Center Utrecht. The study complied with the Declaration of Helsinki. The study protocol was approved by the ethics committees of the UMCU and Maxima Medical Center Eindhoven. All study participants gave written informed consent prior to study inclusion.

Clinical assessment

Comorbidities, current drug use, presence of Raynaud's phenomenon and the presence of visible wounds on hands or wrists were recorded. Subjects filled out a global assessment of disease activity on a Visual Analogue Scale. A swollen and tender joint count was performed by two experienced examiners (KLT and AJLM) in all subjects, in random order. The examiners were blinded to the result of each other's assessments and other study measurements.

Ultrasonography

US was performed by one experienced examiner (DFTC) using a MyLab 60 system (Esaote, Genua, Italy) with an 18-6 MHz linear array transducer. Grey-scale US (GSUS) was performed of metacarpophalangeal (MCP) 1-5 (dorsal side), (P)IP 1-5 of hands (volar side), radiocarpal and midcarpal joints (dorsal side) and flexor and extensor tendons of fingers and wrists.

Power Doppler US (PDUS) was performed only if the GSUS synovitis score was larger than 0. Patient and probe positioning were according to EULAR guidelines (15).

Synovitis on GSUS and PDUS was classified using OMERACT definitions of ultrasonographic joint pathology and graded semi-quantitatively according to a modification of Szkudlarek's grading system (16), combining the criteria for joint effusion and synovial thickening. Since the range of reference values for hypoechoic or anechoic rims around a tendon is very broad, we chose to define tenosynovitis as a hypoechoic or anechoic rim around a tendon exceeding 2 mm (17). This was registered as being absent or present. Grade 1 GSUS synovitis has also been found in healthy subjects (18,19) and is of limited prognostic value in RA (20,21). Therefore, US inflammation was defined as [GSUS synovitis >1 and/or PDUS synovitis >0 and/or GSUS/PDUS tenosynovitis >0]. For individual subjects, the number of joints with inflammation was counted (US joint count). Also, the sum of GSUS synovitis, GSUS tenosynovitis, PDUS synovitis and PDUS tenosynovitis scores were calculated. (US joint index).

Magnetic resonance imaging

Imaging was performed on a 1.5 Tesla MRI. (Philips Ingenia 1,5T, Philips Healthcare, Eindhoven, The Netherlands). The following sequences were acquired: Coronal T1-weighted turbo spin-echo images (TR/TE, 450/7.2; matrix size, 328 x 159; field of view, 18 cm; slice thickness, 2.5 mm), transversal T1-weighted turbo spin-echo images (TR/TE, 500/12; matrix size, 300 x 185; field of view, 12 cm; slice thickness, 2.5 mm), Coronal T2-weighted short-tau inversion recovery (STIR) images (TR/TE, 3251/150; inversion time, 150; matrix size, 328 × 140; field of view, 18 cm; slice thickness, 2.5 mm). T1 images were acquired before and after intravenous gadolinium contrast administration. Images were scored by one experienced rheumatologist (P. Conaghan, MD PhD, University of Leeds, United Kingdom) according to Outcome Measurements in Rheumatology Clinical Trials (OMERACT) Rheumatoid Arthritis MRI Scoring (RAMRIS) method, evaluating synovitis in MCP2-5 and the wrists. The same image set was scored twice with at least a 2-week interval; results were averaged. Similar to US, the number of joints with synovitis was counted (MRI joint count).

9

Optical transmission measurements

OST measurements were performed with the FHP in a standardized way operated by a rheumatology nurse (AJLM). Both hands were inserted through cylindrical openings that contained pressure cuffs. LED lights (wavelengths of 660 and 810 nm) illuminated the (P)IP, MCP and wrist joints of both hands and reference areas from the palmar side. Light transmitted through the joints and reference areas was recorded continu-
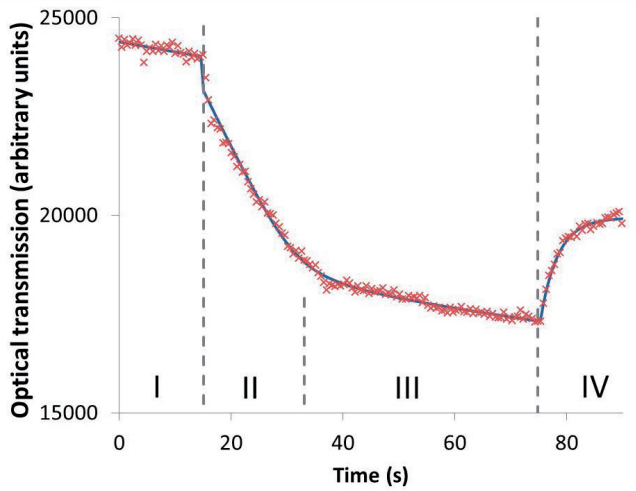
Figure 1. The optical response curve. The recorded intensity of each pixel in a region of interest (ROI) versus the measurement time was fitted to a curve representing the response to inflation and deflation of the pressure cuff. Phase I baseline transmission (inflation of the cuff to 5 mmHg for the duration of 15 seconds), phase II and III inflation of the cuff to 50 mmHg during 60 seconds and phase IV after deflation of the cuff to 5 mmHg during 15 seconds.

ously at the dorsal side by charge coupled device camera with a frame rate of at least 6 per second, alternatingly for the 660 nm wavelength, the 810 nm wavelength and without illumination (background image). A complete measurement was performed within 90 seconds: first, inflation of the cuff to 5 mmHg (15 seconds), second, inflation of the cuff to 50 mmHg (60 seconds) and finally deflation of the cuff (15 seconds).

Image analysis was performed by PBLM using in-house developed software (In-Flame RA120159, December 10, 2012). Regions of interest (ROI) were selected for all joints (joint ROI) and for a position distal to each joint (reference ROI). A ROI consisted of approximately 100 pixels. The recorded intensity of each pixel in a ROI versus the measurement time was fitted to a curve representing the response to inflation and deflation of the pressure cuff, separately for the 660 nm and the 810 nm wavelength (the optical transmission response curve, Figure 1). The optical response curve of a pixel could be described by eight curve parameters per wavelength, i.e. 16 curve parameters for both wavelengths combined. The average of a curve parameter over pixels in a ROI was calculated, giving 16 parameters per ROI. Reference ROIs allowed for correction for systemic effects unrelated to inflammation, such as body temperature and the use of vasoactive medication. This was done by subtracting reference ROI-parameters from joint ROI parameters and then dividing this difference by the average of reference ROI-parameters and joint ROI parameters. Thus, image analysis yielded 16 normalized parameters per joint for use in regression analyses.

Statistical analysis

We developed and internally validated an algorithm for detection of joint inflammation by OST with US as reference by multiple regression analysis. The dependent variable was defined as the maximum of GSUS and PDUS scores for synovitis and tendinitis, and the normalized joint parameters as independent variables. A stepwise forward selection procedure with adjusted R-square testing was used to determine which variable to add. This was repeated until either R-square no longer increased (cut-off value of 0) or, to prevent overfitting of the model, a maximum of four parameters had been selected. This was done separately for each joint region, so separately for (P)IP, MCP and wrists. The regression analysis with the four parameters as independent variables per joint region, was then performed using leave-one-out-cross-validation (leaving out the patient for whom predictions were made after performing linear regression on the data of the other subjects). Then, for each joint in each patient, the linear regression model corresponding to (P)IPs, MCPs or wrists was used to predict joint inflammation. Thus, for each joint in each patient a quantitative score for inflammation was obtained. For individual subjects, an OST joint index over all joints ((P)IP 1-5, MCP 1-5 and wrists of both hands) was calculated. The diagnostic performance of OST was compared with clinical examination, DAS28, US and MRI by receiver operating curve (ROC) analyses, separately at the patient level (one score per patient) and individual joint level (one score per joint). Presence (yes/no) of (teno)synovitis on ultrasonography was used as reference. Difference between areas under the ROC (AUC) was tested for statistical significance using the nonparametric approach developed by DeLong, DeLong, and Clarke-Pearson. Correlations were calculated using Spearman's rank correlation coefficient. Differences between group medians were tested for significance using either the Mann-Whitney U test for 2 groups, or the Kruskal-Wallis 1-way ANOVA for 3 or more groups. P-values below 0.05 (two sided tests) were considered to be statistically significant. The multiple regression analyses were performed using Hemics in-house software (InFlame RA120159), all other analyses by SPSS.

9

RESULTS

Subjects

Demographic and clinical data are reported in Table 1. Table 2 depicts the number and severity of affected PIP, MCP and wrist joints per patient with clinical examinations, ultrasonography and OST.

Table 1. Patient demographic and clinical data.

	Arthralgia	Rheumatoid arthritis		
		$DAS_{28} < 2.6$	$2.6 \geq DAS_{28} \leq 5.1$	$DAS_{28} > 5.1$
Patients (Number)	10	20	26	13
Age (Year)	39±8	48±15	59±8	57±12
Female (%)	100	70	58	54
Duration of RA (Year)	-	3 (2-4)	9 (3-17)	1 (0-6)
ACPA positivity (%)	0	65	77	69
IgM Rheumatoid factor positivity (%)	40	60	77	67
DAS_{28}	3.2±0.9	1.7±0.4	3.8±0.8	5.9±0.6
General Health (VAS)	53 (36-62)	10 (4-37)	51 (32-60)	53 (42-77)
ESR (mm 1st hour)	9 (5-13)	6 (3-10)	14 (8-24)	37 (21-55)
Tender Joints* (Number)	3 (1-7)	0 (0-1)	3 (1-8)	12 (6-16)
Swollen Joints* (Number)	0 (0-0)	0 (0-1)	2 (1-4)	7 (4-10)
Raynaud's phenomenon (%)	0	0	4	15
Use of betablockers (%)	0	5	23	0
Use of calcium channel blockers (%)	0	15	0	0
Use of Conventional DMARDs** (%)	0***	95	92	100
No. of presently used conventional DMARDs	0 (0-0)	1 (1-1)	1 (1-1)	1 (1-2)
Use of methotrexate (%)	0	85	65	69
Use of prednisone (%)	0	0	15	31
Use of biologic DMARDs (%)	0	40	39	8

Numbers are presented as mean ± SD or median (interquartile range) unless mentioned otherwise. * Mean of the two physical examiners. Examined joints: shoulders, elbows, wrists, MCP and (P)IP joints of hands and knees ** methotrexate, sulfasalazine, leflunomide and/or hydroxychloroquine *** One patient used mycophenolic acid for eczema

Table 2. Involvement of wrists, MCP and PIP joints.

	Arthralgia	Rheumatoid Arthritis		
		$DAS_{28} < 2.6$	$2.6 \geq DAS_{28} \leq 5.1$	$DAS_{28} > 5.1$
No. of affected joints				
Tender Joints (0-28)*	1 (0-4)	0 (0-1)	2 (1-6)	11 (5-13)
Swollen Joints (0-28)*	0 (0-0)	0 (0-1)	2 (1-4)	6 (4-9)
GSUS > 0 (0-22)	1 (0-1)	1 (1-3)	4 (2-5)	5 (2-9)
GSUS > 1 (0-22)	0 (0-0)	0 (0-0)	1 (0-2)	2 (1-5)
PDUS > 0 (0-22)	0 (0-0)	0 (0-1)	1 (0-3)	2 (1-6)
GSUS > 0 and/or PDUS>0 (0-22)	1 (0-1)	1 (1-3)	3 (2-5)	5 (2-9)
GSUS > 1 and/or PDUS>0 (0-22)	0 (0-0)	0 (0-1)	1 (0-4)	2 (1-6)
GSUS Tenosynovitis (0-22)	0 (0-0)	0 (0-0)	0 (0-0)	0 (0-0)
PDUS Tenosynovitis (0-22)	0 (0-0)	0 (0-1)	0 (0-2)	1 (0-5)
US inflammation (GSUS >0 and/or PDUS>0 and/or PDUS tenosynovitis) (0-22)	1 (0-1)	2 (1-3)	4 (2-6)	5 (2-11)
US inflammation (GSUS > 1 and/or PDUS>0 and/or PDUS tenosynovitis) (0-22)	0 (0-0)	1 (0-2)	2 (0-5)	3 (2-7)
MRI joint count (synovitis>0) (0-6)	-	5 (3-5)	4 (4-5)	-
MRI joint count (bone marrow oedema) (0-6)	-	1 (0-1)	1 (0-1)	-
Severity of affected joints				
GSUS joint index (synovitis+tenosynovitis) (0-66)	1 (0-1)	2 (1-3)	5 (2-7)	7 (2-13)
PDUS joint index (synovitis+tenosynovitis) (0-66)	0 (0-0)	0 (0-1)	2 (0-5)	3 (1-9)
US joint index (GSUS synovitis + GSUS tenosynovitis + PDUS synovitis + PDUS tenosynovitis) (0-132)	1 (0-1)	2 (1-4)	6 (3-13)	13 (4-27)
RAMRIS synovitis (0-21)	-	6 (4-8)	6 (4-7)	-
RAMRIS bone marrow edema (0-69)	-	1 (0-2)	1 (0-2)	-
RAMRIS erosions (0-230)	-	5 (3-7)	12 (3-17)	-

9

Table 2. (continued)

	Arthralgia	Rheumatoid Arthritis		
		$DAS_{28} < 2.6$	$2.6 \geq DAS_{28} \leq 5.1$	$DAS_{28} > 5.1$
RAMRIS inflammation (synovitis+bone marrow edema) (0-90)		8 (4-10)	6 (6-9)	
RAMRIS (0-320)	-	13 (8-18)	19 (10-23)	-
OST joint index***	4.13 (3.42 – 5.49)	4.40 (3.48-5.10)	5.90 (4.37-7.49)	7.35 (4.27 – 10.28)

Number of affected joints are presented as median (interquartile range)

* Mean of the two physical examiners

** MCP-1-5 and radiocarpal and intracarpal joints on right side only, MRI was performed in patients with $DAS_{28} < 2.6$ at recruitment; all patients with $DAS_{28} < 2.6$ at the study day and additionally 7 patients with $2.6 \geq DAS_{28} \leq 5.1$ (median (IQR) DAS28 of these 7 patients was 3.2 (2.9-3.3)).

*** Kruskal-Wallis 1-way ANOVA p=0.001

Optical transmission

All subjects tolerated the OST measurements well; no adverse events or side effects were observed. Figure 2 depicts an example of the results of OST, US and MRI in a patient in clinical and DAS28 remission. Five subjects had Raynaud's phenomenon and 8 used drugs that may alter peripheral blood flow. Graphically, OST scores of these subjects were not outliers.

Comparison of OST with clinical examination, US and MRI

Median OST was different between the four groups (p=0.001). The correlation coefficients between OST and US and between OST and MRI were higher than between OST and clinical examination (Figure 3).

Joint level: Comparison of OST scores with those at US

Diagnostic performance of OST was tested with inflammation as detected by US as reference standard (Figure 4). Overall, OST had a good performance (AUC 0.83, 95% CI 0.79-0.86).

When testing separately for the (P)IP, MCP and wrists joints, the OST of the (P)IP and MCP joints showed a good discrimination (AUC of 0.79, 95% CI 0.72-0.86,
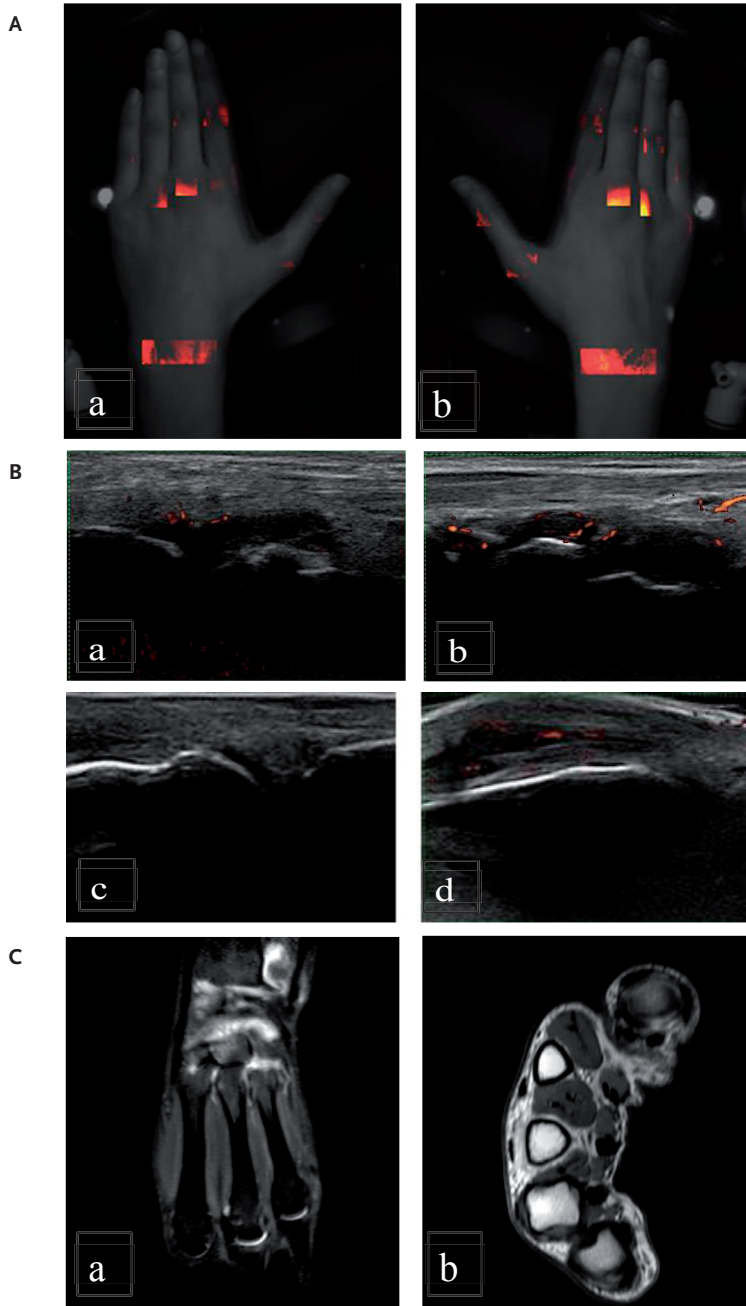
Figure 2. Subclinical inflammation in a patient with rheumatoid arthritis in clinical remission.
Representative images of a patient without clinically detectable arthritis. All imaging techniques show synovitis of both wrists and MCP3 of the right hand. (A) Both hands with Full Hand Proto, (B) GS and PD ultrasonography of joints of left (a) and right wrist (b) and left and right MCP3 (c,d) (C) MRI of right wrist T2 STIR (a) and MCP joints of the right MCP (b, T1 with gadolinium enhancement).

A. OST versus clinical examination

r=0.42
p=0.001

r=0.50
p<0.0001

r=0.25
p=0.037
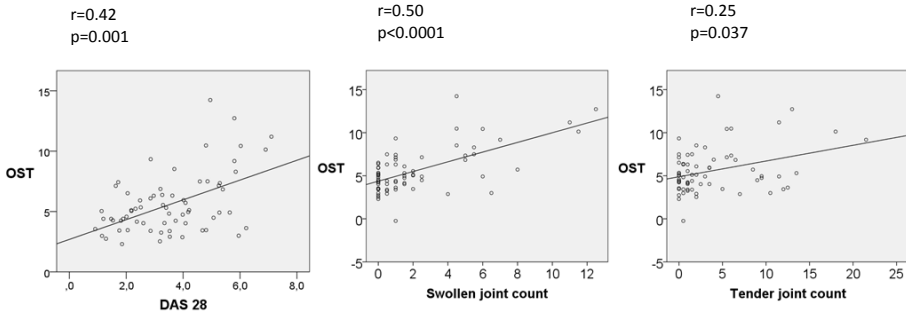
B. OST versus ultrasonography

B1: joint count

r=0.61
p<0.0001

r=0.62
p<0.0001

r=0.64
p<0.0001

B2. Joint index

r=0.57
p<0.0001

r=0.63
p<0.0001

r=0.60
p<0.0001

C. OST versus MRI

r=0.52
p=0.005

r=0.11
p=0.575

r=0.27
p=0.178
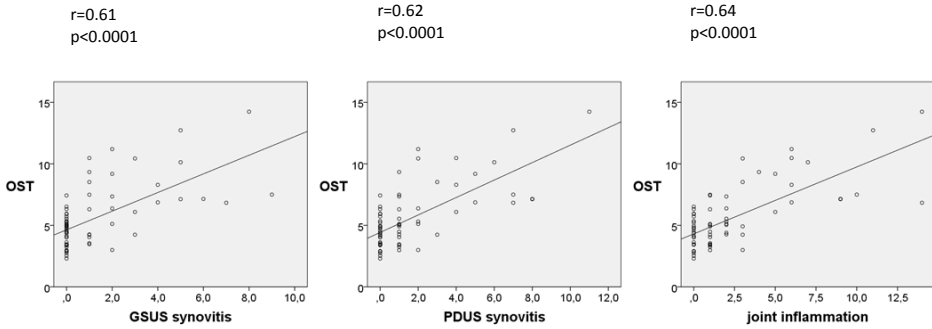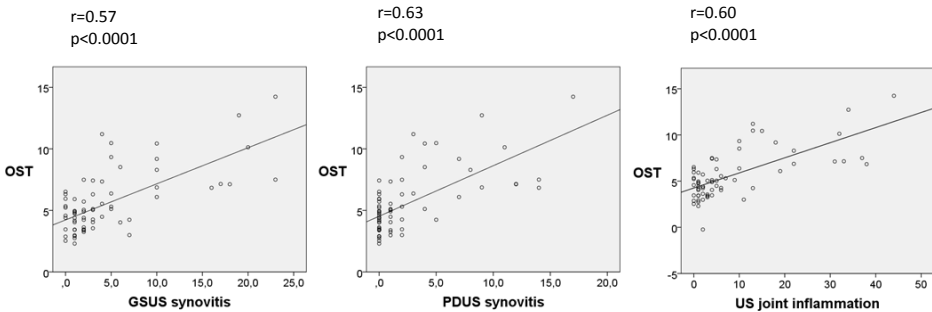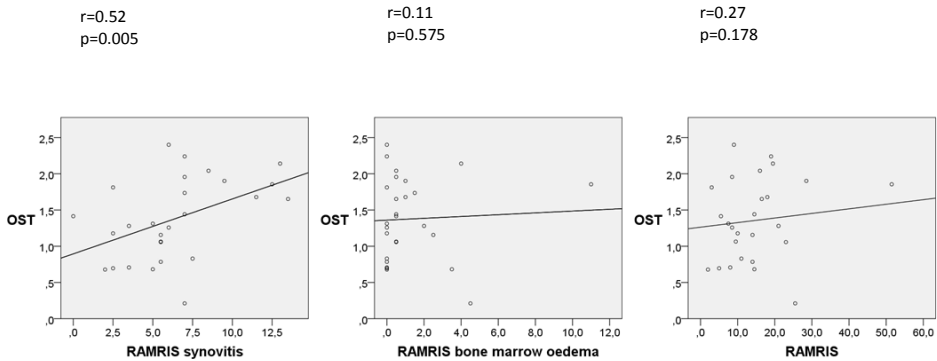
Figure 3. Patient level: relationship between OST and clinical examination, ultrasonography and MRI. A. Correlation between OST and clinical examination (DAS28, swollen joint count of 28 joints and tender joint count of 28 joints). B. Correlation between OST and US, upper three panels with count of number of joints with GSUS synovitis (left panel), count of number of joints with PDUS synovitis (middle panel) and count of joints with inflammation with US (right panel). US inflammation was defined as [GSUS synovitis >1 or PDUS synovitis >0 and/or GSUS /PDUS tenosynovitis >0]. Lower three panels show correlation of OST with joint indexes (sum of semiquantative US scores) of GSUS synovitis (left panel), PDUS synovitis (middle panel) and US inflammation (right panel). C. Correlation between OST and MRI. RAMRIS (left panel) and components of RAMRIS (RAMRIS synovitis (middle panel) and RAMRIS bone marrow oedema (right panel).

p<0.0001 and 0.77, 95% CI 0.71-0.83, p<0.0001, respectively). For the wrists it was poor (AUC 0.64, 95% CI 0.540.74, p=0.006). The values for inflammation as defined by OST with maximum sensitivity and specificity were 0.26 for the MCP-joints (sensitivity of 70%, specificity of 74%), 0.11 for the PIP joints (sensitivity of 83%, specificity of 64%) and 1.0 for the wrists (sensitivity of 39%, specificity of 87%). In comparison, the test characteristics of clinical examination of swollen joints were sensitivity 59%, specificity 86% for the PIP joints, a sensitivity 42% and a specificity of 93% for the MCP joints and a sensitivity 37%, and specificity of 89% for the wrists.

9

OST measurements in clinical decision making

Further, we explored whether OST measurements would be helpful in clinical decision making. We assumed for this that a clinician would want to know if there is joint inflammation in at least one area (reference value: ≥1 joint(s) with joint inflammation on US).

The AUC of OST joint index was 0.75 (95% CI 0.65-0.87, p<0.001) and for the DAS28 0.67 (95% CI 0.53-0.79, p=0.03). This difference was, however, not statistically significant (0.10, 95% CI -0.05 – 0.26, p=0.18).
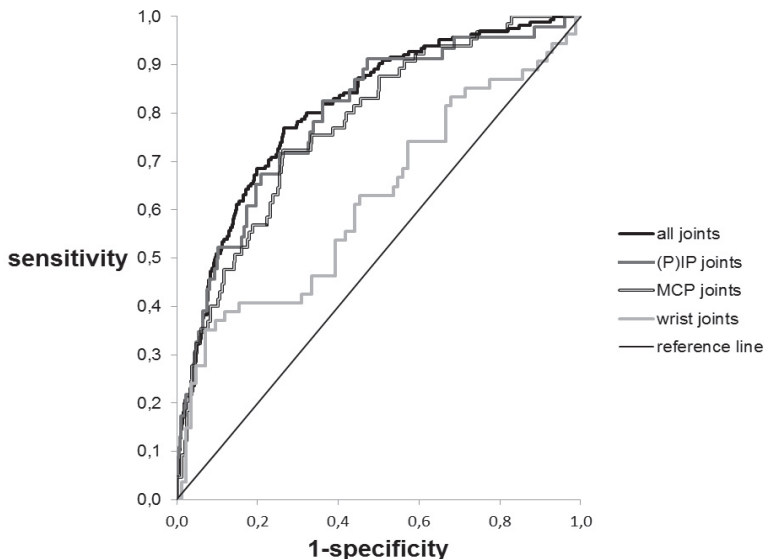
Figure 4. AUC between OST and US at the joint level. Area under receiver operating curve of the OST in all joints (AUC 0.81, 0.77-0.84, p<0.0001), (P)IP joints (AUC 0.79, 0.72-0.86, p<0.0001), MCP joints (AUC 0.78, 0.71-0.83, p<0.0001) and wrists (0.62, 0.52-0.72, p=0.018). (US as reference).

DISCUSSION

In this study, we found that OST measurements agreed well with clinical assessments, US and MRI. The level of agreement was highest with items (directly) reflecting synovial vascularity, such as PDUS and MRI synovitis. As light does not penetrate through bone, we found no agreement between OST and MRI bone marrow edema. Diagnostic performance of OST was better in small hand joints (MCP and PIP joints) than in the wrists, possibly because wrists are larger joints, with less penetration of light. Further, it has been shown that the interobserver reliability of US of the wrists is only moderate (22,23), which may be related to the (more) complex anatomy of the wrist in comparison with PIP and MCP joints. In an exploratory analysis, we tested the value of OST in clinical decision making. OST performed well in detecting joint inflammation (AUC 0.75), however the difference with DAS28 (AUC 0.66) was not statistically significant.

Studies on Fluorescence Optical Imaging (FOI), an invasive technique using light to detect inflammation, have been published earlier (12,13,24-26). It seems that in FOI as well as in OST, the diagnostic performance is inversely related to the size of the joint. The differences in design (such as investigated joints areas, diagnosis) and technique used in the FOI studies hamper the full comparison of these results with our own.

However, our study shows at least equivalent, or even slightly higher sensitivity and specificity at detecting synovitis at the joint level, than the commercially available device for FOI. Additional advantages of our technique are its non-invasiveness and the generation of quantitative results ruling out interobserver variability.

This study has some limitations. There was a risk of model overfitting because US was used both as reference standard in the development of the algorithm and in the evaluation of diagnostic performance. We have tried to reduce overfitting by using a relatively large development sample and leave-one out cross-validation. We think this strategy was successful because we also found a correlation between OST and MRI synovitis and MRI was not used in development of the algorithm. However, validation in a new cohort is still mandatory (27,28). Another possible limitation is that we used subjects with arthralgia as controls. In clinical practice, one wants to differentiate between arthritis and arthralgia. However, it is well known that in a subgroup of subjects, arthralgia may evolve in inflammatory arthritis and subclinical inflammation may already be present (18). None of the subjects in our study developed inflammatory arthritis (> 1 year after study completion). The diagnostic performance of OST in the wrists was rather poor, which may be related to the limited tissue penetration of light through larger structures. It is subject of further studies whether technical adjustments leading to more light penetration through the wrists, are able to improve the accuracy of OST measurements of the wrists. Finally, and similar to other imaging techniques, before OST can be used to monitor patients with rheumatoid arthritis, its sensitivity to change should be investigated. Also, a treat-to-OST-target should be investigated that improves patient outcomes. New studies to evaluate diagnostic performance and sensitivity to change in a new cohort of patients with RA are planned.

In conclusion, OST is a new imaging technique that generates fast, quantitative and operator-independent results. OST performed well in the detection of joint inflammation in patients with RA and is therefore potentially useful in the follow-up of patients with RA.

Note: Abbreviations are explained in the abbreviations list in the appendix.

9

REFERENCES

1   Schoels M, Knevel R, Aletaha D, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search. Ann Rheum Dis 2010;69:638-43.

2   Uhlig T, Kvien TK, Pincus T. Test-retest reliability of disease activity core set measures and indices in rheumatoid arthritis. Ann Rheum Dis 2009;68:972-5.

3   van Gestel AM, Prevoo ML, van 't Hof MA, et al. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. Arthritis Rheum 1996;39:34-40.

4   Ton E, Bakker MF, Verstappen SM, et al. Look beyond the disease activity score of 28 joints (DAS28): tender points influence the DAS28 in patients with rheumatoid arthritis. J Rheumatol 2012;39:22-7.

5   Jacobs JW, Ten Cate DF, van Laar JM. Monitoring of rheumatoid arthritis disease activity in individual patients: still a hurdle when implementing the treat-to-target principle in daily clinical practice. Rheumatology 2014 Published Online First: 28 August 2014.

6   Brown AK, Quinn MA, Karim Z, et al. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. Arthritis Rheum 2006;54:3761-73.

7   Smolen JS, Aletaha D, Bijlsma JW, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 2010;69:631-7.

8   Sinex JE. Pulse oximetry: principles and limitations. Am J Emerg Med 1999;17:59-67.

9   Beuthan J, Zabarylo U, Krause A, et al. RA diagnostics using laser-optical images and conventional X-rays (fused imaging). Med Laser Appl 2007;127-33.

10 Prapavat V, Runge W, Mans J, et al. The development of a finger joint phantom for the optical simulation of early inflammatory rheumatic changes. Biomed Tech 1997;42:31926.

11 Lasker JM, Fong CJ, Ginat DT, et al. Dynamic optical imaging of vascular and metabolic reactivity in rheumatoid joints. J Biomed Opt 2007;12:052001-052001-13.

12 Werner SG, Langer HE, Ohrndorf S, et al. Inflammation assessment in patients with arthritis using a novel in vivo fluorescence optical imaging technology. Ann Rheum Dis 2012;71:504-10.

13 Meier R, Thurmel K, Moog P, et al. Detection of synovitis in the hands of patients with rheumatologic disorders: diagnostic performance of optical imaging in comparison with magnetic resonance imaging. Arthritis Rheum 2012;64:2489-98.

14 Meier AJ, Rensen WH, de Bokx PK, et al. Potential of optical spectral transmission measurements for joint inflammation measurements in rheumatoid arthritis patients. J Biomed Opt 2012 ;17: 081420.

15 Backhaus M, Burmester GR, Gerber T, et al. Guidelines for musculoskeletal ultrasound in rheumatology. Ann Rheum Dis 2001;60:641-9.

16 Szkudlarek M, Court-Payen, Jacobsen S,et al. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. Arthritis Rheum 2003;48:955-62.

17 Schmidt WA, Schmidt H, Schicke B, et al. Standard reference values for musculoskeletal ultrasonography. Ann Rheum Dis 2004;63:988-94.

18 van de Stadt LA, Bos WH, Meursinge RM, et al. The value of ultrasonography in predicting arthritis in auto-antibody positive arthralgia patients: a prospective cohort study. Arthritis Res Ther 2010;12: R98.

19 Balsa A, de Miguel E, Castillo C, et al. Superiority of SDAI over DAS-28 in assessment of remission in rheumatoid arthritis patients using power Doppler ultrasonography as a gold standard. Rheumatology 2010;49:683-90.

20 Gartner M, Mandl P, Radner H,et al. Sonographic joint assessment in rheumatoid arthritis: associations with clinical joint assessment during a state of remission. Arthritis Rheum 2013;65:2005-14.

21 Witt M, Mueller F, Nigg A, et al. Relevance of grade 1 gray-scale ultrasound findings in wrists and small joints to the assessment of subclinical synovitis in rheumatoid arthritis. Arthritis Rheum 2013; 65:1694-701.

22 Ohrndorf S, Fischer IU, Kellner H et al. Reliability of the novel 7-joint ultrasound score: results from an interand intraobserver study performed by rheumatologists. Arthritis Care Res 2012;64:1238-43.

23 Dopazo GN, Ten Cate DF, Swen WA, et al. The most reliable probe position in the ultrasonographic examination of the wrist in rheumatoid arthritis. Clin Exp Rheumatol 2012;30:566-9.

24 Meier R, Thuermel K, Noel PB, et al. Synovitis in patients with early inflammatory arthritis monitored with quantitative analysis of dynamic contrast-enhanced optical imaging and MR imaging. Radiology 2014;270:176-85.

25 Schafer VS, Hartung W, Hoffstetter P, et al. Quantitative assessment of synovitis in patients with rheumatoid arthritis using fluorescence optical imaging. Arthritis Res Ther 2013;15:R124.

26 Werner SG, Langer HE, Schott P, et al. Indocyanine green-enhanced fluorescence optical imaging in patients with early and very early arthritis: a comparative study with magnetic resonance imaging. Arthritis Rheum 2013;65:3036-44.

27 Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15: 361-87.

28 Altman DG, Vergouwe Y, Royston P, at al. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605.

9

CHAPTER 10
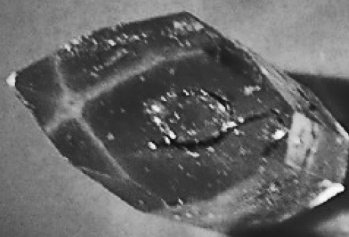
General discussion

The heterogeneity of rheumatoid arthritis (RA), the difficulty of diagnosing early RA and of assessing disease activity were the rationales to investigate the use of ultrasound (US) for clinical purposes in RA. This thesis was divided into two domains: a fundamental domain and a clinical domain.

The aims of this thesis were:
1. To increase the validity of greyscale ultrasound (GSUS) and power Doppler ultrasound (PDUS)
2. To increase the reliability of GSUS and PDUS for wrists
3. To evaluate the added, predictive value of US in patients with active RA

Main findings:
1. Current definitions for GSUS may be too sensitive for clinical purposes, when scanning for synovitis and tenosynovitis (chapter 2 and chapter 3). The PDUS modality is highly variable between machines (chapter 4)
2. The most reliable probe position in scanning the wrist joint longitudinally is the line from Lister's tubercle to the third digit (chapter 5)
3. Standard US has no added value in predicting clinical effect at one year in RA patients treated to target (chapter 8)

We have seen in the introduction of this thesis that US became a feasible and accepted tool in rheumatology. However, there are several good reasons to investigate the role of US specifically in RA further. These rationales need to be viewed in the light of the paradigmatic change in the treatment of RA patients, i.e. 'treat to target' and 'tight control' (1, 2).

Some unresolved issues remained however, that could be divided into two domains, fundamental and clinical. In the fundamental domain, the main issues were with validity and reliability. In the clinical domain, the main issue concerns the added value of US in predicting clinical outcomes of RA patients treated to target.

FUNDAMENTAL DOMAIN -- MAIN ISSUES

10

GS synovitis has been defined by the OMERACT based on the echogenicity of the tissue (3). This is a qualitative definition. It has been tried to make this more quantitative by creating semiquantitative scoring systems. These usually range from 0 to 3, with 0 indicating no US synovitis and 3 indicating severe US synovitis (3-7). However, the validity of GSUS has already been questioned by the observation that common definitions for GSUS inflammation may be too sensitive when evaluated in RA patients

and non-arthritic controls (8-11). In up to 88% of non-arthritic controls a greyscale score of at least one, when using a semiquantitative system, may be found. This is an important observation, because it may hold essential implications for decisions regarding escalation of therapy. For instance, if the target of the treatment is absence of synovitis, false positive synovitis scoring could lead to unnecessary therapeutic actions. So a further validation of GSUS was warranted. Although validation of PDUS with histological examination had been done in the past (12-14), GSUS validation with fundamental modalities such as histology or anatomical dissection was still lacking.

Therefore, the often seen anechoic to hypoechoic area on the dorsal aspect of proximal phalanges of fingers (**chapter 2**), which could potentially jeopardize current scoring systems, deserved further investigation. We showed that this area, coined the 'distal anechogenicity in the metacarpophalangeal joint' (DAEM), is not a sign of (subclinical) arthritis and not a US artefact. This is because we found the DAEM in a large proportion of healthy subjects who did not develop arthritis, and we found the structure at anatomical dissection of the joint. Our findings might contribute to the high prevalence of mild synovitis scores in non-arthritic individuals and healthy controls (8-11). The relevance of minimal GSUS scores in the absence of PDUS have been questioned by others (15, 16). This suggests that the interpretation of US scores need re-evaluation. Zufferey et al. suggested to use a cumulative GSUS score of at least 8 as a cut off when scanning 22 joints (16). This cut off value needs to be investigated in larger trials, possibly even in histological validation studies. In our view, until then, we propose to regard GSUS score 1 as false positive, unless there is also power Doppler signal present in the joint.

In **chapter 3** we investigated the preoperative prevalence of (subclinical) US tenosynovitis using a binary score based on the echogenicity around the tendon, following the definition that was used in those days (3). We compared this to peroperative examination and postoperative histological examination in a cohort of idiopathic CTS patients, who underwent carpal tunnel release surgery. We found histological tenosynovitis to be rare, although tenosynovitis as defined on GSUS and peroperatively by the surgeon was frequently present. In these patients reactive changes were found histologically, but not histological inflammation. This was in line with the total absence of PD signal in any of the wrists. Possibly, the abnormalities found at US and by the surgeon peroperatively were a response of the tissue to mechanical stressors or pressure in the carpal tunnel (17). The low prevalence of tenosynovitis in our study was similar to the prevalence found in other studies (18, 19).

Our findings could have clinical implications as they may question the use of glucocorticoid injections in these patients. More importantly, from a 'validation' point of view, the definitions of US tenosynovitis apparently result in false positive scores with histology as a reference standard. After we were finished collecting the data for

our study, a new semiquantitative tenosynovitis scoring system was published, with 0 indicating no tenosynovitis and 3 indicating severe tenosynovitis (20). Analogous to the semiquantitative scoringsystems for synovitis this is based on subjective criteria for abnormality. We could not re-evaluate our findings using this new scoring system since US is a dynamic technique. However, looking at the pictures these authors present in their paper, we think that similar issues regarding false positivity may arise because of the grade 1 tenosynovitis, analogous to the grade 1 synovitis in chapter 2. It is unfortunate that the publication of our study has been at a similar time as this paper, since this paper states that no surgical or histological validation of US scores for tenosynovitis exist. In our opinion, their semiquantitative scoring system could have profited from the findings in our study. Until a new semiquantitative scoring system is published, we propose to focus primarily on PDUS when assessing tenosynovitis, in line with what the authors of the semiquantitative scoringsystem suggest (20).

As mentioned before, PD proved to be a very valuable US modality in MSK US. It enables detection of very low flows in very small vessels, necessary to detect the capillaries formed in the neovascularization of inflammation. Therefore, PDUS has the potential to display (subclinical) joint and tendon inflammation early. This is exactly what is needed in RA as it would pave the way for earlier detection of disease activity and more precise assessment of disease activity in RA patients. To investigate the validity of this statement we built a flow phantom and assessed the PD modality of 5 US machines, of which we presented the results in **chapter 4**. In this study we showed that the sensitivity of the PD modalities of 5 US machines was very different. The differing sensitivities were not the only finding. We also found that some of the machines could not even detect flow velocities that were sufficiently low to detect flow in capillaries indicating subclinical inflammation. The consequence of this finding is that the absence of PD signal does not necessarily mean absence of inflammation.

Flow phantoms had been published before (12, 21-23), however, the size of the capillaries used was either larger than one would expect in an inflamed joint, or one could not be certain of the flow in the individual capillaries, making it difficult to compare the findings in those studies with our own. Either way, older machines were used, which considerably limits comparing the results of our study with the older studies as well. However, in our study we made assumptions on capillary sizes and flow velocities in vivo. These assumptions were not based on studies of blood flow in inflamed joints, but on studies on blood flow in periulcerous regions. Therefore, more work is needed to ascertain flows and capillary sizes in inflamed joints which is crucial for studying early inflammation in RA.

On a different note, this study showed that difficulties may arise in research where you least expect them. In conducting this study we experienced problems in creating the blood mimicking fluid (BMF). It should have similar US characteristics as blood,

10

but at the same time, it should not block the capillaries. This appeared to be quite difficult especially in the smallest capillaries; after much trial and error we managed to find a suitable composition with which we could execute our study.

Different sensitivities to low flows on different US machines means a patient could be designated as having US inflammation, i.e. possible escalation of medication on one machine, whereas the same patient could be designated as being free of inflammation on the other machine, i.e. possible tapering of medication. This is a very unwelcome situation. Our suggestion is that for future studies, as well as in clinical practice, US machines should be tested on a flow phantom similar to ours. Even more so, US machine manufacturers should test their machines on a reference standard such as our flow phantom, with small capillaries and low flows, in the development process. This has not been effectuated yet.

The last major issue in the fundamental domain is the issue of reliability, which is the topic of **chapter 5**. We conducted a study to find the most reliable probe position to scan the wrist. Three ultrasonographers used three probe positions to scan RA patients' wrists longitudinally. We found that the probe position defined by the line between Lister's tubercle and digit III was the most reliable position to scan the wrist, as compared to the position from Lister's tubercle to digit II and the ulnocarpal position. As we have already seen in the introduction of this thesis, reliability remains an issue in rheumatological US. Besides variation introduced by different examiners, the reliability statistic itself is equivocal. The measures used to calculate reliability are sensitive to study sample heterogeneity. This means that studies with very high or very low prevalence of the outcome is, the reliability measures tend to be an underestimation of reliability. Additionally, for ICC's specifically, the equation used influences the outcome. Therefore, we started chapter 5 by introducing these main issues regarding the ICC and a possible approach to curtail them. We propose to have an independent expert select cases for the reliability exercise to increase the heterogeneity in the study sample. Also, we found that calibration improved the reliability in our study, a concept that had been described for MRI specifically (24, 25). The most important proposition on our side was, however, that when using the ICC for expression of the reliability, one should always define the precise method which was used to calculate the ICC, since we know that ICCs can range from 0.17 to 0.91 depending on the equation used (26). The reliability in our wrist study remained rather modest, which was, at least partly, due to the above mentioned, low heterogeneity of the study sample. Since the current measures for reliability are limited due to its dependency on the number and distribution of abnormalities scored (27), new studies, possibly from a theoretical, statistical point of view are warranted. Despite all this, the results of our study suggest to use the position defined as the line from Lister's tubercle to digit III when scanning the wrist longitudinally in daily, clinical practice and trials.

CLINICAL DOMAIN -- MAIN ISSUES

Part 2 of this thesis, the clinical considerations of US in RA, begins with **chapter 6**, a short chapter on the possible issues with the validity of current clinical measures to assess disease activity. One of the rationales to write this thesis was to investigate the possible added, predictive value of US in patients with active RA. However, current outcome measures are clinical measures. These measures are validated for use on a group level, and prove to be not sufficient to assess disease activity in individual patients (28, 29). Other drawbacks of these measures are underestimation of the true disease activity, e.g. by an insensitive clinical exam, or overestimation of the true disease activity, e.g. in case of concomitant pain syndrome or osteoarthritis (28, 29). Our proposition was that adding US to clinical measures would improve evaluation of disease activity in individual patients. We proceed part 2 with a systematic review to investigate the current role of US in diagnosing RA and evaluation of remission of RA (**chapter 7**). One of the issues in applying US in daily clinical practice is the number of joints needed to scan. In the busy daily, clinical practice as few joints as possible should be scanned. Our review indicates that at least bilateral MCP, wrist and MTP joints should be scanned in aiding diagnosis of RA. For remission, scanning MCPs and wrists of the dominant hand appears to be sufficient. Both in diagnosis and remission, PDUS appears to be more valuable than GSUS. Using GSUS appears to lead to false positive findings of synovitis. This may relate to GSUS scores occurring in healthy controls and non-arthritic individuals as we saw in chapter 2 (8-11, 30). So in this chapter we show that US has added predictive value in UA patients and also in RA patients already in remission (15, 31-35).

The added predictive value of US in newly diagnosed RA patients has not been investigated in depth before. We present the results of the study investigating this question in **chapter 8**. The small joints of hands and feet and the wrists of 174 newly diagnosed RA patients were scanned using US. All patients were treated to target, and were followed using a tight control approach. We investigated whether US improves prediction of persistent active disease (defined as disease activity score (DAS)28>2.6, i.e. no remission) when added to predictors, such as baseline disease activity (36), baseline erosions (37) and presence of auto-antibodies (38). We found that a standard baseline US exam did not improve prediction of persistent active disease. Baseline disease activity, RF positivity and strictly monitoring patients using 4 week intervals (compared to 3 months intervals) were the only predictors of persistent active disease. So following patients every month, aiming for remission, appears to lead to better clinical outcomes than following patients less frequently, or aiming for low disease activity, when controlled for DAS28 at baseline and RF positivity. This suggests that there is still room for improvement, even in the treat to target era using clinical mea-

10

sures. Radiographic progression was too rare to be used as an outcome measure in longitudinal analyses. One earlier paper hinted at the possibility that a US exam holds predictive value (39). In this study, a 'time integrated value' of PDUS correlated to clinical outcome measures at one year. A 'time integrated value' was defined as the area under curve for the respective parameter at baseline, three months and also at one year. Since this type of analysis uses the value for one year also, it is not a demonstration of predictive value at a certain time point. In this study, the baseline US exam did not have predictive value.

The absence of predictive US value raises questions. One could argue that the results are caused by problems with the internal validity of the study as we used a multi-centre design with differing treatment and monitoring strategies. However, this is a reflection of everyday clinical practice in rheumatology. Examiners were well trained and machines did not differ between the centres which reduced the heterogeneity resulting from US exams. Part of the predictive value of US in our study may have been diluted by the fact that all patients were tightly controlled, treated to target (40). Apparently, in these current treatment paradigms a large part of the study sample responds well enough to the treatment so there is not much room left for US to predict clinical outcomes.

Another important finding of this study was that of the 158 patients of whom a follow up US examination was available, 16 patients (10%) did not have ultrasonographic inflammation at baseline. Of these, 13 (81%) did not have ultrasonographic inflammation at baseline or one year follow up. Since all patients included in the study were RA patients according to the ACR(1987) classification criteria this highlights the issues described earlier in the thesis with the clinical measures. It may mean that there is a subset of persons with symptoms in the absence of US inflammation, possibly even in the absence of any inflammation. This deserves further investigation.

Although we found that a standard baseline US exam did not predict clinical outcomes, US may be a very useful adjunct in patients in whom applying the DAS28 may be less valid (28, 29), e.g. with concomitant pain syndrome or obesity. We proposed therefore that new trials be conducted using US as an outcome measure, instead of current clinical measures. In our opinion, it deserves further attention to investigate to what extent US can aid in diagnosing RA and evaluating remission in a prospective cohort with US as an outcome measure. The first initiatives in this field have been currently set in motion (29, 41).

The final chapter (**chapter 9**) describes a study in which a new imaging modality, optical spectral transmission (OST), is evaluated as a measure to detect joint inflammation. A similar technique using OST has been published (42). However, this technique is invasive and not operator independent, as the technique described in our study is. We describe a technique which assesses the venous blood flow in the

hand joints as a proxy for inflammation, by evaluating the amount of light that can be transmitted through the hands. This information is analysed using an algorithm resulting in an operator independent cut off depicting inflammation. US inflammation was used as the reference standard in this study. We found that OST is superior in detecting joint inflammation as compared to clinical examination, even reclassifying patients in clinical remission to being not in remission. Of course, a limitation of this study was that US was used as a reference standard in the development of the OST algorithm, and as an outcome measure when comparing OST and clinical examination. There is a risk of overfitting; however, we used a relatively large development sample and cross validated the model using the leave one out principle. Nonetheless, our findings should be validated in a new study sample. This limitation aside, by using US as a reference standard we paved the way for the development of newer, easy and operator independent measures for detecting disease activity. In our opinion, OST deserves to be investigated further, by conducting studies using a treat to OST-target strategy.

GENERAL METHODOLOGICAL CONSIDERATIONS

Some comments on specific methodological concepts should be made.

Validity & reliability

Validity of US plays an important role in using US in arthritis. It refers to the extent to which the abnormality as seen on US is actually present, and if present, the extent to which the abnormality represents inflammation. Anatomical dissection or histological examination would the reference standard for this. However, assessing each and every joint against histology is hampered by the invasive biopsy that damages joint tissue. The damage dissecting a joint would do is evident. Validity was the subject of our research as described in chapter 2, 3 and 4. We are one of the first to investigate the validity of MSK US on such a fundamental level.

Reliability is the concept that describes the overall consistency of a measure taking into account systemic errors, patient variance and observer variance. The latter plays an important role in US image acquisition as a slightly different probe position may result in a dramatically different image. This may result in different scoring of inflammation between observers. To reduce this source of interobserver variability, fixed anatomical locations were used holding the probe in the same position.

10

Nonetheless, we observed a moderate ICC(A,1) between observers when absolute agreement was high, as described in chapter 5 and 8. The reason for this may be that the ICC very much depends on variance between patients, as we see from the following equation: ICC = Variance (patients) / Variance (patients)+Variance (observer)+Variance (error). We see that the ICC is influenced by the heterogeneity of the study sample; in our studies the distribution of the numbers of inflamed joints. A low variance between patients (or joints) will result in a low ICC, possibly even if absolute agreement is high. Another issue is that ICCs can be calculated using at least 6 different equations (26, 43), with very different results. The conditions that need to apply to use each respective equation are well described. For the reader to appreciate the reported reliability, when expressed as ICC, it is necessary that it has been reported which equation was used.

Patient selection

For the research described in chapter 8, patient selection could have influenced the validity of the results. All patients were participating in trials that had a tight control approach. This approach results in many patients reaching the treatment target and therefore may have diminished the predictive effect of US in this population. Although 'tight control' should also be employed in daily clinical practice, this may not be possible in all patients, who are more heterogeneous than those included in clinical trials, fulfilling selection criteria. In these patients in daily clinical practice who are not always controlled very tightly, US might actually have added predictive value.

NEW INSIGHTS

- Current definitions for GSUS may be too sensitive for clinical purposes, when scanning for synovitis and tenosynovitis (chapter 2 and chapter 3)
- The PDUS modality is highly variable between machines (chapter 4)
- The most reliable probe position in scanning the wrist joint longitudinally is the line from Lister's tubercle to the third digit (chapter 5)
- Standard US has no added value in predicting clinical measures at one year in RA patients treated to target (chapter 8)

RECOMMENDATIONS FOR CLINICAL PRACTICE

The new insights from this thesis have resulted in 5 recommendations for clinical practice.

1. To get acquainted with the anatomical variation of US findings, it is advised, for novice ultrasonographers, to scan healthy subjects first.(Chapter 2, 3)
2. GSUS score 1 should be regarded as a false positive unless a PD signal is present in the synovium. (Chapter 2)
3. Tenosynovitis as scored on GSUS should not necessarily be regarded as inflammation. (Chapter 3)
4. When acquiring a new US machine make sure the PD modality on this specific system is sensitive enough to detect subclinical inflammation. (Chapter 4)
5. It might be considered to use US as a diagnostic tool in patients in whom the clinical examination may be less valid, or in case of doubt. (Chapter 7, 8)

RECOMMENDATIONS FOR FUTURE RESEARCH

If there is one thing I have learned these past years is that US research in rheumatology needs collaboration with other disciplines. Not only medical doctors, such as rheumatologists and surgeons; we need engineers, anatomists and statisticians just as much.

In this light, we should begin with improving the validity of both GSUS and PDUS even further using anatomic specimen and histological samples as a reference standard. This will improve the validity, but will also improve reliability. Using knowledge of the studies improving the validity of US should result in new definitions and new scoring systems. These should be used to perform new reliability studies. Simultaneously, a firm collaboration with statistical departments should be sought. The issues of reliability statistics themselves should be solved.

To be able to detect inflammation as early as possible, preferably in the subclinical situation, even before the window of opportunity, it is of the utmost importance to accelerate the technological improvements of US machines. By focusing on the developments in 3D US, both the validity and the reliability of US exams may improve, even without technological improvements of current US machines. At the same time, it should be attempted to increase the sensitivity of US machines. In this respect the developments in the field of plane wave imaging and contrast imaging are very promising (44, 45).

10

However, when reaching these goals of improving US machines, making them more sensitive, new validity and reliability studies should be executed. Since improving US machines will be an ongoing process it will probably be necessary to continue to execute validity and reliability studies in MSK US.

All this is needed to execute new trials using US as an outcome measure. In other words; we should conduct trials with a 'treat-to-US-target'. This will be a study with a similar design as the one described in chapter 8 of this thesis, only then with US outcomes. In doing this it may be possible to distinguish those patients needing more intensive treatment from the moment of diagnosis, from the patients needing less intensive treatment, paving the path for truly personalized medicine.

In the introduction of this thesis I have discussed the separate entities *diagnostic* US and *therapeutic* US. This entire thesis was on the diagnostic use of US, which is the conventional research domain. However, with the ever improving sensitivity of US machines the time has come 'to travel the road less travelled by' rheumatologists. The first initiatives are already afoot. Using US it is possible to take biopsies directly from the inflamed tissue (46), and carpal tunnel release can be achieved through a percutaneous, sonographically guided release (47, 48). This technique could be investigated in a sample similar to the one described in chapter 3, or even in RA patients. It might not only lead to better outcomes, due to the less invasive nature of the procedure, it is also a field that needs to be looked upon from a cost effectiveness perspective.

Finally, in other fields than rheumatology, there are developments in the field of therapy using contrast bubbles. A distant prospect perhaps, but in theory, contrast bubbles may be coated with antibodies to pathological T-cells, which can then be made porous by US so medication can enter the cell, or be destroyed using US (49). This would possibly decrease the necessary dosage of the drug, thereby decreasing side effects and also costs.

Note: Abbreviations are explained in the abbreviations list in the appendix.

REFERENCES

1. Bakker MF, Jacobs JW, Welsing PM, Vreugdenhil SA, van Booma-Frankfort C, Linn-Rasker SP, et al. Early clinical response to treatment predicts 5-year outcome in RA patients: follow-up results from the CAMERA study. Ann Rheum Dis. 2011 Jun;70(6):1099-103.

2. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. Lancet. 2004 Jul 17-23;364(9430):263-9.

3. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostino MA, et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol. 2005 Dec;32(12): 2485-7.

4. Backhaus M, Ohrndorf S, Kellner H, Strunk J, Backhaus TM, Hartung W, et al. Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. Arthritis Rheum. 2009 Sep 15;61(9):1194-201.

5. Naredo E, Bonilla G, Gamero F, Uson J, Carmona L, Laffon A. Assessment of inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with grey scale and power Doppler ultrasonography. Ann Rheum Dis. 2005 Mar;64(3):375-81.

6. Scheel AK, Hermann KG, Kahler E, Pasewaldt D, Fritz J, Hamm B, et al. A novel ultrasonographic synovitis scoring system suitable for analyzing finger joint inflammation in rheumatoid arthritis. Arthritis Rheum. 2005 Mar;52(3):733-43.

7. Szkudlarek M, Court-Payen M, Jacobsen S, Klarlund M, Thomsen HS, Ostergaard M. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. Arthritis Rheum. 2003 Apr;48(4):955-62.

8. Balsa A, de Miguel E, Castillo C, Peiteado D, Martin-Mola E. Superiority of SDAI over DAS-28 in assessment of remission in rheumatoid arthritis patients using power Doppler ultrasonography as a gold standard. Rheumatology (Oxford). 2010 Apr;49(4):683-90.

9. van de Stadt LA, Bos WH, Meursinge Reynders M, Wieringa H, Turkstra F, van der Laken CJ, et al. The value of ultrasonography in predicting arthritis in auto-antibody positive arthralgia patients: a prospective cohort study. Arthritis Res Ther. 2010;12(3):R98.

10. Witt M, Mueller F, Nigg A, Reindl C, Leipe J, Proft F, et al. Relevance of grade 1 gray-scale ultrasound findings in wrists and small joints to the assessment of subclinical synovitis in rheumatoid arthritis. Arthritis Rheum. 2013 Jul;65(7):1694-701.

11. Guillen Astete CA, Boteanu A, Zea Mendoza A. Comparison of Prevalence of Synovitis by Ultrasound Assessment in Subjects Exposed or Not to Self-Reported Physical Overexertion: The Monday's Synovitis. ScientificWorldJournal. 2014;2014:563981.

12. Koski JM, Saarakkala S, Helle M, Hakulinen U, Heikkinen JO, Hermunen H. Power Doppler ultrasonography and synovitis: correlating ultrasound imaging with histopathological findings and evaluating the performance of ultrasound equipments. Ann Rheum Dis. 2006 Dec;65(12):1590-5.

13. Walther M, Harms H, Krenn V, Radke S, Faehndrich TP, Gohlke F. Correlation of power Doppler sonography with vascularity of the synovial tissue of the knee joint in patients with osteoarthritis and rheumatoid arthritis. Arthritis Rheum. 2001 Feb;44(2):331-8.

14. Walther M, Harms H, Krenn V, Radke S, Kirschner S, Gohlke F. Synovial tissue of the hip at power Doppler US: correlation between vascularity and power Doppler US signal. Radiology. 2002 Oct; 225(1):225-31.

10

15. Peluso G, Michelutti A, Bosello S, Gremese E, Tolusso B, Ferraccioli G. Clinical and ultrasonographic remission determines different chances of relapse in early and long standing rheumatoid arthritis. Ann Rheum Dis. 2011 Jan;70(1):172-5.

16. Zufferey P, Moller B, Brulhart L, Tamborrini G, Scherer A, Finckh A, et al. Persistence of ultrasound synovitis in patients with rheumatoid arthritis fulfilling the DAS28 and/or the new ACR/EULAR RA remission definitions: Results of an observational cohort study. Joint Bone Spine. 2014 Oct;81(5): 426-32.

17. Ettema AM, Amadio PC, Zhao C, Wold LE, An KN. A histological and immunohistochemical study of the subsynovial connective tissue in idiopathic carpal tunnel syndrome. J Bone Joint Surg Am. 2004 Jul;86-A(7):1458-66.

18. Uchiyama S, Itsubo T, Nakamura K, Kato H, Yasutomi T, Momose T. Current concepts of carpal tunnel syndrome: pathophysiology, treatment, and evaluation. J Orthop Sci. 2010 Jan;15(1):1-13.

19. Nakamichi K, Tachibana S. Histology of the transverse carpal ligament and flexor tenosynovium in idiopathic carpal tunnel syndrome. J Hand Surg Am. 1998 Nov;23(6):1015-24.

20. Naredo E, D'Agostino MA, Wakefield RJ, Moller I, Balint PV, Filippucci E, et al. Reliability of a consensus-based ultrasound score for tenosynovitis in rheumatoid arthritis. Ann Rheum Dis. 2013 Aug;72(8):1328-34.

21. Browne JE, Watson AJ, Hoskins PR, Elliott AT. Validation of a sensitivity performance index test protocol and evaluation of colour Doppler sensitivity for a range of ultrasound scanners. Ultrasound Med Biol. 2004 Nov;30(11):1475-83.

22. Kamishima T, Nishida M, Horie T, Narita A, Sagawa A, Henmi M, et al. Power Doppler signal calibration using capillary phantom for pannus vascularity in rheumatoid finger joint. A pilot study. Clin Exp Rheumatol. 2011 Nov-Dec;29(6):1057.

23. Veltmann C, Lohmaier S, Schlosser T, Shai S, Ehlgen A, Pohl C, et al. On the design of a capillary flow phantom for the evaluation of ultrasound contrast agents at very low flow velocities. Ultrasound Med Biol. 2002 May;28(5):625-34.

24. Lukas C, Braun J, van der Heijde D, Hermann KG, Rudwaleit M, Ostergaard M, et al. Scoring inflammatory activity of the spine by magnetic resonance imaging in ankylosing spondylitis: a multireader experiment. J Rheumatol. 2007 Apr;34(4):862-70.

25. Bird P, Joshua F, Lassere M, Shnier R, Edmonds J. Training and calibration improve inter-reader reliability of joint damage assessment using magnetic resonance image scoring and computerized erosion volume measurement. J Rheumatol. 2005 Aug;32(8):1452-8.

26. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979 Mar;86(2):420-8.

27. Ten Cate DF, Luime JJ, Hazes JM, Jacobs JW, Landewe R. Does the intraclass correlation coefficient always reliably express reliability? Comment on the article by Cheung et al. Arthritis Care Res (Hoboken). 2010 Sep;62(9):1357-8; author reply 8.

28. Jacobs JW, Ten Cate DF, van Laar JM. Monitoring of rheumatoid arthritis disease activity in individual patients: still a hurdle when implementing the treat-to-target principle in daily clinical practice. Rheumatology (Oxford). 2014 Aug 28.

29. Wakefield RJ, D'Agostino MA, Naredo E, Buch MH, Iagnocco A, Terslev L, et al. After treat-to-target: can a targeted ultrasound initiative improve RA outcomes? Postgrad Med J. 2012 Aug;88(1042): 482-6.

30. Ten Cate DF, Luime JJ, Hazes JM, Kleinrensink GJ, Jacobs JW. Is the frequent sonographic anechoic area distally in metacarpophalangeal joints a sign of arthritis? Ultrasound Med Biol. 2014 Oct; 40(10):2537-41.

31.   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 Oct;58(10):2958-67.

32.   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 Mar;70(3):500-7.

33.   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 Jan;64(1):67-76.

34.   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 Sep-Oct; 28(5):686-94.

35.   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 Sep;48(9):1092-7.

36.   Romao VC, Canhao H, Fonseca JE. Old drugs, old problems: where do we stand in prediction of rheumatoid arthritis responsiveness to methotrexate and other synthetic DMARDs? BMC Med. 2013;11:17.

37.   Combe B, Dougados M, Goupille P, Cantagrel A, Eliaou JF, Sibilia J, et al. Prognostic factors for radiographic damage in early rheumatoid arthritis: a multiparameter prospective study. Arthritis Rheum. 2001 Aug;44(8):1736-43.

38.   Aletaha D, Alasti F, Smolen JS. Rheumatoid factor determines structural progression of rheumatoid arthritis dependent and independent of disease activity. Ann Rheum Dis. 2013 Jun;72(6):875-80.

39.   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 Feb 15;57(1):116-24.

40.   Quinn MA, Emery P. Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. Clin Exp Rheumatol. 2003 Sep-Oct;21(5 Suppl 31):S154-7.

41.   Dale J, Purves D, McConnachie A, McInnes I, Porter D. Tightening up? Impact of musculoskeletal ultrasound disease activity assessment on early rheumatoid arthritis patients treated using a treat to target strategy. Arthritis Care Res (Hoboken). 2014 Jan;66(1):19-26.

42.   Werner SG, Langer HE, Schott P, Bahner M, Schwenke C, Lind-Albrecht G, et al. Indocyanine green-enhanced fluorescence optical imaging in patients with early and very early arthritis: a comparative study with magnetic resonance imaging. Arthritis Rheum. 2013 Dec;65(12):3036-44.

43.   McGraw K, Wong S. Forming inferences about some intraclass correlation coefficients. Psycological Methods 1996;1(1):30-46.

44.   Tanter M, Fink M. Ultrafast imaging in biomedical ultrasound. IEEE Trans Ultrason Ferroelectr Freq Control. 2014 Jan;61(1):102-19.

45.   Klauser AS, Franz M, Bellmann Weiler R, Gruber J, Hartig F, Mur E, et al. Contrast-enhanced ultrasonography for the detection of joint vascularity in arthritis--subjective grading versus computer-aided objective quantification. Ultraschall Med. 2011 Dec;32 Suppl 2:E31-7.

46.   Scire CA, Epis O, Codullo V, Humby F, Morbini P, Manzo A, et al. Immunohistological assessment of the synovial tissue in small joints in rheumatoid arthritis: validation of a minimally invasive ultrasound-guided synovial biopsy procedure. Arthritis Res Ther. 2007;9(5):R101.

10

47. Schreiber AL, Sucher BM, Nazarian LN. Two novel nonsurgical treatments of carpal tunnel syndrome. Phys Med Rehabil Clin N Am. 2014 May;25(2):249-64.

48. McShane JM, Slaff S, Gold JE, Nazarian LN. Sonographically guided percutaneous needle release of the carpal tunnel for treatment of carpal tunnel syndrome: preliminary report. J Ultrasound Med. 2012 Sep;31(9):1341-9.

49. Kooiman K, Foppen-Harteveld M, van der Steen AF, de Jong N. Sonoporation of endothelial cells by vibrating targeted microbubbles. J Control Release. 2011 Aug 25;154(1):35-41.
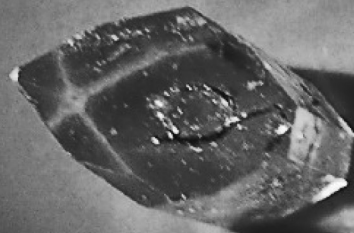
English and Dutch summary

Rheumatoid arthritis (RA) is a common disease but should be considered more a clinical syndrome than one distinct disease. Early RA is difficult to diagnose because clear signs and symptoms may be absent in this stage. Furthermore, the assessment of disease activity using clinical measures may not be valid enough, overestimating or underestimating true disease activity in certain patients. For these reasons, ultrasonography (US) has been introduced in RA; for more objective diagnosis and disease assessment. US is a technique that uses high frequent sound waves to create an image. This image is displayed in grey hues (grey scale ultrasound (GSUS)), indicating different tissues, or in color (power Doppler ultrasound (PDUS)), displaying vessels. Despite the progress ultrasonographic research has made in RA, some important issues remained. These issues were in the fundamental domain, i.e. validity and reliability, and in the clinical domain, especially the added predictive value of US in RA patients. The goal of this thesis was to address these issues.

The aims of this thesis were:
1. To increase the validity of greyscale ultrasound (GSUS) and power Doppler ultrasound (PDUS)
2. To increase the reliability of GSUS and PDUS for wrists
3. To evaluate the added, predictive value of US in patients with active RA

Chapter 1, the introduction, of this thesis provides an overview of the rationale to use US in RA, by discussing relevant events and developments in the history of RA and US. Also, relevant concepts of the technique of US are addressed. The remaining issues with the use of US in RA are discussed and an outline of this thesis is presented. This thesis is divided in two domains, a fundamental domain and a clinical domain.

FUNDAMENTAL DOMAIN

The first issue that we dealt with was the issue of validity. In **chapter 2**, we investigate the frequently occurring anechoic to hypoechoic area on the dorsal aspect of proximal phalanges of fingers, coined the 'distal anechogenicity in the metacarpophalangeal joint' (DAEM). This area could potentially jeopardize current scoring systems. We evaluated the prevalence of the DAEM in 24 non-arthritic subjects and compared its dimensions in 10 non-arthritic subjects and 7 RA patients. The DAEM is not a sign of (subclinical) arthritis nor a US artifact. This is the conclusion, because we found the DAEM in a large number of the non-arthritic subjects who did not develop arthritis, and the structure proved to be an extension of the joint space at anatomical dissection

11

of the joint. In our view, we propose to regard GSUS score 1 as too sensitive for clinical purposes, unless there is also power Doppler signal present in the same joint.

The validity of definitions for US tenosynovitis were discussed in **chapter 3**, where we investigated the preoperative prevalence of (subclinical) US tenosynovitis. This was compared to peroperative examination and postoperative histological examination in a cohort of idiopathic CTS patients. We found histological tenosynovitis to be rare, although tenosynovitis as defined on GSUS and peroperatively by the surgeon was frequently present. Our findings indicate that current definitions of US tenosynovitis may result in false positive scores with histology as a reference standard. We propose to focus primarily on PDUS when assessing tenosynovitis.

In **chapter 4** we investigate the validity of the PD modality by comparing the sensitivity of the PD modalities of 5 US machines. We showed the machines performed very differently. This means a patient could be designated as having US inflammation according to PD on one machine, whereas the same patient could be designated as being free of inflammation on the other machine. This is a very unwelcome situation. Our suggestion was that in future studies, and also when testing machines for clinical practice, US machines should be tested on a flow phantom similar to ours. Manufacturers should test their machines on a reference standard such as our flow phantom, with small capillaries and low flows, in the development process. This has not been effectuated yet.

The last major issue in the fundamental domain is the issue of reliability, which is the topic of **chapter 5**. We begin with a short introduction on the issues with reliability statistics, especially the intraclass correlation coefficient. The main issue is that with very high or very low prevalence of the outcome, reliability measures tend to underestimate reliability. Next we describe a study to find the most reliable probe position to scan the wrist. Three ultrasonographers used three probe positions to scan RA patients' wrists longitudinally. We found that the probe position defined by the line between Lister's tubercle and digit III was the most reliable position to scan the wrist, as compared to the position from Lister's tubercle to digit II and the ulnocarpal position. We propose to use this most reliable position in daily, clinical practice and trials.

CLINICAL DOMAIN

Part 2 of this thesis, the clinical considerations of US in RA, begins with **chapter 6**, a short chapter on the possible issues with the validity of current clinical measures to assess disease activity. These measures are validated for use on a group level, and prove to be not sufficient to assess disease activity in individual RA patients. Our

proposition was that adding US to clinical measures might improve evaluation of disease activity in individual patients.

In **chapter 7** we present a systematic review to investigate the current role of US in diagnosing RA and evaluation of remission of RA. Our review indicates that at least bilateral metacarpophalangeal (MCP), wrist and metatarsophalangeal (MTP) joints should be scanned in aiding diagnosis of RA. For remission, scanning MCPs and wrists of the dominant hand only appears to be sufficient. Both in diagnosis and remission, PDUS appears to be more valuable than GSUS. Using GSUS appears to lead to false positive findings of synovitis, as in controls also often GSUS signs or arthritis are found. In this chapter we show that US has added predictive value in undifferentiated arthritis (UA) patients and in RA patients in remission.

In **chapter 8** we present the results of the study investigating the added value of US in predicting clinical outcomes in newly diagnosed RA patients. The small joints of hands and feet and the wrists were scanned using US. All patients were treated to target, and were followed using a tight control approach. A standard baseline US exam did not improve prediction of persistent active disease. Baseline disease activity, RF positivity and strictly monitoring patients using 4 week intervals (compared to 3 months intervals) were the only predictors of persistent active disease. This last predictor suggests that there is still room for improvement, even in the treat to target era using clinical measures, since patient's are usually followed using intervals longer than 4 weeks. An important finding of this study was that a subset of the patients did not have ultrasonographic inflammation at baseline or at one year follow up. Since all patients were RA patients according to the ACR(1987) classification criteria this highlights the issues described earlier in the thesis with the clinical measures. It may mean that there is a subset of persons with symptoms in the absence of US inflammation, possibly even in the absence of any inflammation. This deserves further investigation.

In our opinion, it deserves further attention to investigate prospectively to what extent US measures can be used as target for treatment, or using US as basis for treatment escalation and tapering. The first initiatives in this field have been currently set in motion.

The final chapter (**chapter 9**) describes a study in which a new imaging modality, optical spectral transmission (OST), is evaluated as a measure to detect joint inflammation. This is assessed by evaluating the amount of light that can be transmitted through the hands with and without an inflated blood pressure cuff. We found that OST is superior in detecting joint inflammation as compared to clinical examination. Although our findings should be validated in a new study sample, in conducting this study, we paved the way for the development of newer, easy and operator independent measures for detecting disease activity. In our opinion, OST deserves to be investigated further, by performing studies using a treat to OST-target strategy.

To go back to the aims of this thesis:

1. To increase the validity of greyscale ultrasound (GSUS) and power Doppler ultrasound (PDUS)
2. To increase the reliability of GSUS and PDUS for wrists
3. To evaluate the added, predictive value of US in patients with active RA

We showed that the validity of both GSUS and PDUS can be increased further. Regarding reliability we showed that reliability statistics may not be entirely valid; the most reliable position to scan the wrist is defined by the line between Lister's tubercle and digit III. We found that a baseline US examination did not improve prediction of clinical outcomes in early RA patients treated to target.

Reumatoïde artritis (RA) komt relatief veel voor, maar kan beter gezien worden als een klinisch syndroom dan een scherp af te grenzen ziekte-entiteit. Ook is RA in een heel vroege fase moeilijk te diagnosticeren doordat dan duidelijke tekenen en symptomen vaak nog ontbreken. En als de diagnose gesteld is, blijkt het bepalen van de ziekteactiviteit van de individuele patiënt met de huidige maten niet valide genoeg, met overschatting van de ziekteactiviteit bij sommige patiënten en onderschatting ervan bij andere. Onder meer om een objectievere maat te hebben voor het bepalen van de ziekteactiviteit is men ultrasonografie (US), of echografie, gaan gebruiken bij RA. US is een techniek die gebruik maakt van hoogfrequente geluidsgolven om een afbeelding te maken. Deze afbeelding bestaat uit de kleur zwart en vooral verschillende tinten grijs (Engels: grey scale ultrasound (GSUS)), die respectevelijk vocht en verschillende soorten weefsel weergeven, of uit kleuren (Engels: (power Doppler ultrasound (PDUS)), die bewegende structuren, vrijwel altijd bloed en bloedvaten, weergeven. Ondanks dat er vooruitgang is geboekt binnen het echo-onderzoek bij RA, blijven bepaalde kwesties op basaal niveau, dat van validiteit en betrouwbaarheid, als op klinisch niveau, bijvoorbeeld in hoeverre US voorspellende waarde heeft voor RA, onopgelost. Het doel van dit proefschrift was om een aantal van deze vraagstukken op te lossen.

De globale doelen van dit proefschrift waren:
1. Het vergroten van de validiteit van GSUS en PDUS.
2. Het vergroten van de betrouwbaarheid van echografisch onderzoek naar artritis van polsen.
3. Het bepalen van de toegevoegde, voorspellende waarde van echografie bij patiënten met actieve RA.

Hoofdstuk 1, de inleiding van dit proefschrift geeft een overzicht van belangrijkste redenen om US te gebruiken bij RA, aan de hand van de geschiedenis van RA en US. Verder worden in de inleiding enkele essentiële concepten van de techniek van US uit de doeken gedaan. Aan het eind van dit hoofdstuk worden nog actuele, basale en klinische kwesties van US bij RA geïntroduceerd. Het proefschrift wordt vervolgens dienovereenkomstig ingedeeld in twee domeinen, een basaal domein en een klinisch domein.

11

BASAAL DOMEIN

De eerste kwestie die we behandelen betreft validiteit. In **hoofdstuk 2**, onderzoeken we de vaak voorkomende anechoïsche tot hypoechoïsche structuur aan de dorsale zijde van de proximale phalanx van vingers, door ons de 'distale anechogeniciteit in het metacarpophalangeaal gewricht' (DAEM) gedoopt. De DAEM zou heel gevoelig scoren op artritis volgens huidige scoringssystemen op losse schroeven kunnen zetten. We onderzochten de prevalentie van de DAEM in 24 proefpersonen zonder artritis en vergeleken de afmetingen van de DAEM bij 10 proefpersonen zonder artritis en bij 7 RA patiënten. De DAEM bleek geen teken te zijn van subklinische artritis, noch een echografisch artefact. We concludeerden dit, omdat we de DAEM in een groot aantal proefpersonen zonder artritis, die evenmin artritis ontwikkelden, hebben aangetoond. Bij anatomische dissectie bleek het een omslagplooi van de gewrichtsholte te zijn. Hierom, en omdat bij mensen zonder artritis vaak een GSUS score 1 van een gewricht gevonden wordt, stellen wij dat de GSUS score 1 te gevoelig is om in de kliniek te gebruiken, tenzij er ook power Doppler signaal in hetzelfde gewricht aanwezig is.

De validiteit van het echografisch vaststellen van tenosynovitis wordt in **hoofdstuk 3** onderzocht, waarin we preoperatief de prevalentie van echografische tekenen van (subklinische) tenosynovitis onderzochten. We vergeleken deze met de peroperatieve beoordeling op eventuele tenosynovitis door de chirurg en het postoperatieve histologische onderzoek op tenosynovitis, in een cohort van patiënten met idiopathische carpale tunnelsyndroom (CTS). Bij histologisch onderzoek bleek tenosynovitis zelden voor te komen, terwijl het echografisch en peroperatief vaker aanwezig was. Dit suggereert dat de huidige echografische definities van tenosynovitis fout-positieve scores kunnen opleveren, als histologie de gouden standaard is. Ons voorstel is om voor het echografisch diagnosticeren van tenosynovitis vooral waarde te hechten aan PDUS.

In **hoofdstuk 4** wordt de validiteit van de PDUS modaliteit onderzocht door het vergelijken van de gevoeligheid van PDUS van 5 verschillende types echomachines. We laten in dit hoofdstuk zien dat de gevoeligheid van de afzonderlijke machines erg verschilt. Dit zou kunnen betekenen dat bij een patiënt met een bepaald echo-apparaat wél echografisch tekenen van ontsteking worden aangetoond en met een ander echo-apparaat niet. Dit is een bijzonder ongewenste situatie. We stelden voor om echo-apparaten voor toekomstig wetenschappelijk onderzoek, maar ook voor klinische doeleinden te testen op een flowfantoom, zoals dat wij gebouwd hebben. Bestaande apparaten kunnen dan zo goed mogelijk ingesteld kunnen worden en fabrikanten zouden in het ontwikkelingsproces van nieuwe apparaten ook hun voordeel kunnen doen met ons flowfantoom met kleine capillairen en heel lage flows.

De laatste kwestie in het basale domein, die in **hoofdstuk 5** behandeld wordt, betreft betrouwbaarheid. Dit hoofdstuk begint met een korte inleiding over knelpunten

die er ten aanzien van het bepalen van betrouwbaarheid zijn, vooral betreffende de 'intraclass' correlatie coëfficiënt. Het belangrijkste probleem is dat als de uitkomst heel vaak, of juist zelden voorkomt, de betrouwbaarheid wordt onderschat. Na deze introductie volgt de beschrijving van een onderzoek, waarin we de meest betrouwbare probe-positie bij het scannen van het polsgewricht op artritis onderzoeken. Drie echografisten hebben drie longitudinale probe-posities getest bij het scannen van de pols. Het bleek dat de meest betrouwbare probe-positie ligt op de denkbeeldige lijn tussen het tuberkel van Lister en de basis van de derde vinger; deze positie werd vergeleken met de positie gedefinieerd door een lijn van het tuberkel van Lister naar de basis van de tweede vinger en met een ulnocarpale positie. Ons voorstel is dat zowel in de dagelijkse praktijk als in onderzoeksverband deze meest betrouwbare probe-positie wordt gebruikt.

KLINISCH DOMEIN

Deel 2 van dit proefschrift behandelt klinische kwesties en begint met **hoofdstuk 6**, een kort hoofdstuk, waarin validiteit van huidige klinische ziekteactiviteitsmaten onder de loep wordt genomen. Deze maten zijn gevalideerd op groepsniveau en blijken niet te voldoen om de ziekte activiteit van individuele RA patiënten betrouwbaar weer te geven. Ons voorstel is om echografie aan klinische maten toe te voegen, om de ziekte activiteit van individuele patiënten nauwkeuriger te bepalen.

In **hoofdstuk 7** beschrijft een systematische review de rol van echografie bij het diagnosticeren van RA en de plaatsbepaling van echografie bij RA patiënten in remissie. De resultaten van onze review suggereren dat tenminste beiderzijds metacarpophalangeaal (MCP), polsen en metatarsophalangeaal (MTP) gewrichten gescand moeten worden als hulpmiddel bij het diagnosticeren van RA. Bij RA patiënten klinisch in remissie lijkt het toereikend om slechts de MCPs en de pols aan dominante zijde te scannen. Zowel voor diagnose als remissie lijkt PDUS waardevoller te zijn dan GSUS. Enkel GSUS gebruiken leidt tot fout-positieve scores, gezien het feit dat bij individuen zonder inflammatoire gewrichtsaandoeningen ook vaak lage GSUS-scores van inflammatie vastgesteld worden. Onze review laat zien dat echografie additief voorspellende waarde heeft in de groep kort bestaande, nog ongedifferentieerde artritispatiënten en bij RA patiënten klinisch in remissie.

In **hoofdstuk 8** worden de resultaten gepresenteerd van een onderzoek naar de toegevoegde waarde van echografie bij het voorspellen van klinische uitkomsten van recent gediagnosticeerde RA-patiënten. Hand-, polsen kleine voetgewrichten werden op baseline echografisch onderzocht. Alle patiënten werden behandeld met het

11

therapeutisch doel lage ziekteactiviteit of remissie. Het routinematige echografische onderzoek bij diagnose blijkt geen toegevoegde voorspellende waarde te hebben voor het voorspellen van klinisch persisterend actieve ziekte na een jaar behandelen. Ziekte-activiteit ten tijde van diagnose, reumafactortest-positiviteit en frequent monitoren en in medicatie bijsturen van patiënten (4 wekelijks versus 3 maandelijks) waren wel voorspellers van persisterende ziekteactiviteit. De laatste voorspeller suggereert dat er nog ruimte voor verbetering is in de klinische praktijk, aangezien het interval waarmee ook patiënten met kort bestaande RA poliklinisch gecontroleerd worden, vaak langer is dan 4 weken. Een andere belangrijke bevinding van dit onderzoek is dat een aantal patiënten echografisch geen gewrichtsontsteking had ten tijde van de klinische diagnose en evenmin na 1 jaar follow-up. Aangezien alle patiënten voldeden aan de ACR(1987) classificatie criteria voor RA benadrukt deze bevinding de kwesties die in hoofdstuk 6 aan de orde zijn. Dit is iets dat nader onderzoek verdient.

Volgens ons moet er nog beter onderzocht worden in hoeverre echografie op indicatie kan worden gebruikt als uitkomstmaat bij behandeling, en of echografie kan worden gebruikt bij het intensiveren of juist afbouwen van behandeling. De eerste stappen in deze zijn reeds gezet.

Het laatste hoofdstuk (**hoofdstuk 9**) beschrijft een onderzoek waarin een nieuwe beeldvormende modaliteit, optische spectraal transmissie (OST), is onderzocht om gewrichtsontsteking te meten. Deze techniek gebruikt de mate van transmissie van licht door handen met en zonder opgeblazen bloeddrukband aan de onderarm om ontsteking van polsen vingergewrichten in maat en getal uit te drukken. Het lijkt erop dat OST even goed en misschien beter dan lichamelijk onderzoek gewrichtsontsteking zou kunnen detecteren. Ondanks dat deze bevindingen in een nieuw onderzoek moeten worden gevalideerd, bijvoorbeeld in trials waarin RA patiënten worden behandeld op basis van OST uitkomsten, hebben we de weg geëffend voor de ontwikkeling van een nieuwe, makkelijk te gebruiken en onderzoeker-onafhankelijke methode om ziekteactiviteit te meten. .

Samenvattend, teruggrijpend op de doelen van dit proefschrift:
1. Het vergroten van de validiteit van GSUS en PDUS.
2. Het vergroten van de betrouwbaarheid van echografisch onderzoek naar artritis van polsen.
3. Het vaststellen wat de toegevoegde, voorspellende waarde van echografie in patiënten met actieve RA is.

In dit proefschrift is beschreven dat de validiteit van zowel GSUS als PDUS verbeterd kan worden. Ten aanzien van betrouwbaarheid hebben we laten zien dat betrouwbaarheidsmaten zelf matig valide kunnen zijn, en dat de meest betrouwbare probe-positie

om de pols op artritis te scannen op de lijn van het tuberkel van Lister naar de basis van de derde vinger ligt. Tenslotte, met behulp van een routinematig echografisch onderzoek ten tijde van de diagnose RA wordt, indien de patiënten intensief behandeld worden, de voorspelling van het klinische effect na een jaar behandelen niet beter.

11
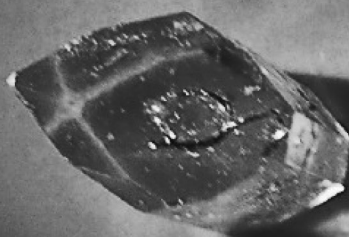
APPENDIX

APPENDIX A: LIST OF ABBREVIATIONS

2-D	two dimensional ultrasound
3-D	three dimensional ultrasound
A	amplitude (Pascal)
aCCP	anti-cyclic citrullinated peptide
ACR	American College of Rheumatology
ANOVA	analysis of variance
ASA	American Society of Anaesthesiologists
AUC	area under the receiver-operating curve
BCE	before the common era
BMF	Blood-mimicking fluid
c	velocity of the sound wave (meter/second)
CAMERA	computer-assisted management in early RA
CD	color Doppler
CI	confidence interval
cm/s	centimeter per second
CRP	C-reactive protein
CTS	Carpal tunnel syndrome
DAEM	distal anechogenicity in the metacarpophalangeal joint
DAS	disease activity score
DAS28	disease activity score assessing in 28 joints
DIP	distal interphalangeal;
DMARD	disease modifying antirheumatic drug
DRU	distal radio-ulnar
ESR	erythrocyte sedimentation rate
EULAR	European league against rheumatism
F	female
f	frequency (Hertz(Hz))
FHP	Full Hand Prototype
FM	fibromyalgia
FOI	Fluorescence Optical Imaging
GS	greyscale
GSUS	greyscale ultrasonography/ultrasound
HAQ	health assessment questionnaire
HLA-B27	human leukocyte antigen B27
Hz	Hertz
IA	inflammatory arthritis
IC	intercarpal
ICC	intraclass correlation coefficient
IC-joint	intercarpal joint

A

IQR	interquartile range
JE	joint effusion
kHz	kilo Hertz
λ	wavelength (meter)
LED	light emitting diode
LR	likelihood ratio test
LSRA	long standing RA
μm	micro-meter
M	male
MCP	metacarpophalangeal
MDRD	modification of diet in renal disease
MHz	Mega Hertz
mm²	square mm
mmHg	millimetres of mercury
mo	month
MRI	magnetic resonance imaging
MSK	musculoskeletal
MT	midtarsal
MTP	metatarsophalangeal
MTX	methotrexate
NA	Non applicable
Nm	newton metre
NSAID	non steroid anti-inflammatory drug
OMERACT	Outcome Measures in Rheumatology
OR	odds ratio
OST	optical spectral transmission
PD	power Doppler
PDUS	power Doppler ultrasonography/ultrasound
PFA	paraformaldehyde
PIP	proximal interphalangeal
PRF	pulse repetition frequency
Q	flow (m³/second)
R	radius (m)
RA	rheumatoid arthritis
RAMRIS	rheumatoid arthritis MRI synovitis score
RC	radiocarpal
Ref.	reference
RF	rheumatoid factor
ROC	receiver operating characteristic
ROI	region of interest
ROMA	reumatologie online monitoring applicatie
RSTL	radial styloid

SD	standard deviation
SDAI	simplified disease activity index
SDD	smallest detectable difference
SH	synovial hypertrophy;
Sj	swollen joints
SJC	swollen joint count
SONAR	sound navigation and ranging
STIR	short-tau inversion recovery
SvH	Sharp-van der Heijde
TCZ	Tocilizumab
TJC	tender joint count
tREACH	treatment in the Rotterdam early arthritis cohort
TS	tenosynovitis
UA	undifferentiated arthritis
UAE	U Act Early
UC	ulnocarpal
UMCU	university medical center Utrecht
US	ultrasonography/ultrasound
USJC	ultrasound joint count
USTL	ulnar styloid
VAS	visual analogue scale
V_{avg}	average flow velocity (m/second)
w/w	weight per weight
y.	year

A

APPENDIX B: LIST OF FIGURES

Chapter 1

Figure 1: http://spie.org/Images/Graphics/Newsroom/Import-ed-2011/003998/003998_10_fig1.jpg

Figure 2: http://www.ob-ultrasound.net/project/kraut3.jpg

Figure 3: http://en.wikipedia.org/wiki/Ultrasonic_testing#mediaviewer/File:-UT_principe.svg

Figure 4: From the author's personal collection

Figure 5: From the author's collection

Figure 6: From the author's collection

Figure 7: http://www.daviddarling.info/childrens_encyclopedia/-Ultrasound_For_Kids.html

Figure 8: Provided by Esaote

Figure 9: http://www.ob-ultrasound.net/project/phased.jpg

Figure 10: http://cfile27.uf.tistory.com/image/263E3B38539BF343212420

Figure 11: Frinking PJ, Bouakaz A, Kirkhorn J, Ten Cate FJ, de Jong N, Ultrasound Med Biol. 2000;26:965-75

A

APPENDIX C: SUPPLEMENTAL FILES

Chapter 3

Additional file 1. US versus surgical definition of tenosynovitis in 33 wrists

		Surgery		
		TS+	TS-	
Ultrasonography	TS+	16	3	20
	TS-	13	1	14
		29	4	33

TS + is defined as tenosynovitis grade 1 and grade 2 combined (see box 1)

Additional file 2. US versus histological definition of tenosynovitis in the 27 wrists from which biopsies were taken

		Histology		
		TS+	TS-	
Ultrasonography	TS+	4	11	16
	TS-	1	11	12
		5	23	27

For histology, TS + is defined as grade 2 tenosynovitis (see box 2)

Chapter 7

Additional file 1. Search strategies

Early RA

Pubmed: (ultrasonography[mesh] OR ultrasonography[sh] OR ultrasono*[tw] OR echogr*[tw]) AND (synovi*[tw] OR arthrit*[tw] OR inflammatory joint condition*[tw] OR joint inflammation*[ti] OR joint inflammation*[tw] OR oligoarthrit*[tw] OR polyarthrit*[tw]) AND ((index*[tw]) OR (score*[tw]) OR (scori*[tw]) OR (count[tw]))

Embase: (inflammat* NEAR/3 joint*):de,ab,ti OR oligoarthrit*:de,ab,ti OR polyarthrit*:de,ab,ti OR 'arthritis'/exp OR arthritis OR synovi* OR arthrit*:de,ab,ti AND (ultrasono* OR echogr*) AND (index*:de,ab,ti OR score*:de,ab,ti OR scori*:de,ab,ti OR count*:de,ab,ti)

Number of articles in search

A
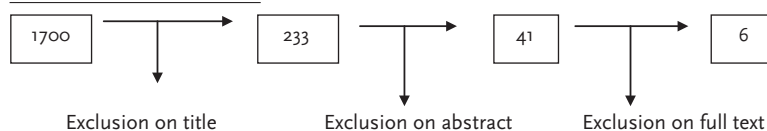
Remission

Pubmed: (ultrasonography[mesh] OR ultrasonography[sh] OR ultrasono*[tw] OR echogr*[tw]) AND (arthrit*[tw] OR inflammatory joint condition*[tw] OR joint inflammation*[ti] OR joint inflammation*[tw] OR oligoarthrit*[tw] OR polyarthrit*[tw] OR disease activity[tw] OR DAS) AND remission

Embase: (arthrit*:ti,ab,de OR (inflammat* NEAR/3 joint*):ti,ab,de OR oligoarthrit*:ti,ab,de OR polyarthrit*:ti,ab,de) OR ('disease activity':ti,ab,de) OR (DAS) OR (arthritis/exp OR arthrit*:ti,ab,de) AND (ultrasono* OR echogr*) AND remission

Number of articles in search

| 387 | | 23 | | 16 | | 11 |

Exclusion on title   Exclusion on abstract   Exclusion on full text

---
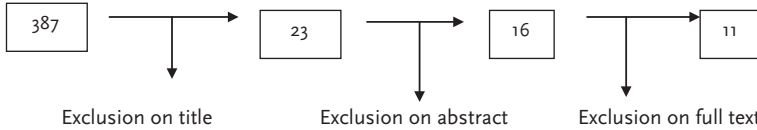
Additional file 2

Box 1: Adaptation of the phases (levels) of Sackett & Haynes of diagnostic questions in diagnostic studies

- Phase 1. Do signs of US inflammation in patients with the condition differ from those without the condition?
- Phase 2. Are patients with certain signs of US inflammation more likely to have the condition?
- Phase 3. Do signs of US inflammation distinguish patients with and without the condition among those in whom it is clinically sensible to suspect the condition?
- Phase 4. Do patients undergoing the US fare better in their ultimate health outcome than similar untested patients?

For the evaluation of studies on diagnosing RA, the condition of level 1, 2 and 3 is the diagnosis RA. For studies evaluating remission, this condition is the presence of clinical remission, according to several definitions.

Additional file 3. Quality assessment lists
Quadas-2, with extra question on sample size.
DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection:

Q1A Was a consecutive or random sample of patients enrolled?	Yes/No/Unclear
Q1B Was a case-control design avoided?	Yes/No/Unclear
Q1C Did the study avoid inappropriate exclusions?	Yes/No/Unclear
Q1D Was the sample size appropriate?	Yes/No/Unclear
Q1E Could the selection of patients have introduced bias?	**RISK: LOW/HIGH/UNCLEAR**

B. Concerns regarding applicability

Describe included patients (prior testing, presentation, intended use of index test and setting):

Q1F Is there concern that the included patients do not match the review question?	**CONCERN: LOW/HIGH/UNCLEAR**

DOMAIN 2: INDEX TEST(S)
If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted:

Q2A Were the index test results interpreted without knowledge of the results of the reference standard?	Yes/No/Unclear
Q2B If a threshold was used, was it pre-specified?	Yes/No/Unclear
Q2C Could the conduct or interpretation of the index test have introduced bias?	**RISK: LOW /HIGH/UNCLEAR**

B. Concerns regarding applicability

Q2D Is there concern that the index test, its conduct, or interpretation differ from the review question?	**CONCERN: LOW /HIGH/UNCLEAR**

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted:

Q3A Is the reference standard likely to correctly classify the target condition?	Yes/No/Unclear
Q3B Were the reference standard results interpreted without knowledge of the results of the index test?	Yes/No/Unclear
Q3C Could the reference standard, its conduct, or its interpretation have introduced bias?	**RISK: LOW /HIGH/UNCLEAR**

B. Concerns regarding applicability

Q3D Is there concern that the target condition as defined by the reference standard does not match the review question?	**CONCERN: LOW /HIGH/UNCLEAR**

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram):
Describe the time interval and any interventions between index test(s) and reference standard:

Q4A Was there an appropriate interval between index test(s) and reference standard?	Yes/No/Unclear
Q4B Did all patients receive a reference standard?	Yes/No/Unclear
Q4C Did patients receive the same reference standard?	Yes/No/Unclear

A

Q4D Were all patients included in the analysis?	Yes/No/Unclear
Q4E Could the patient flow have introduced bias?	**RISK: LOW /HIGH/UNCLEAR**

Quality assessment remission:

Quality assessment✣

Population

	Item
a	Number of patients
b	Age, (mean/median +- SD/range)
c	Sex; (mean/median +- SD/range)
d	time since diagnosis; (mean/median +- SD/range)
e	time since remission; (mean/median +- SD/range)
f	therapy used at moment of remission
g	medication changes during period of remission
h	DAS(28) (mean/median +- SD/range)
i	Baseline erosions
j	HAQ(mean/median +- SD/range)
k	RF/aCCP
l	ESR/CRP

Study design

	Item
m	Longitudinal(n?)
n	Blinded
o	US/CR/CE in short time period at each timepoint
p	definition diagnosis
q	definition remission
r	definition of erosive disease
s	Definition flare (if applicable)
t	currently known covariates recorded

| Analysis	
	Item
u	Multivariate logistic regression
v	Number of cases >=10 per covariate
w	RR/OR +- 95%CI
x	Definitions US inflammation clear

* The quality of each parameter is described as "Good" when present, "Bad" when absent or "Unclear" when it is not clear if the parameter is present.

Chapter 8

Additional file: Univariate and multivariate logistic regression in the multiple imputation dataset (M=10) using STATA 12 and REALCOM

Variable	Univariate		Multivariate	
	OR (95% CI)	p-value	OR (95% CI)	p-value
DAS28 per point	1.31 (1.00 – 1.69)	0.04	1.53 (1.12 – 2.08)	0.007
Female gender	1.84 (0.91-3.72)	0.09	1.94 (0.76-2.86)	0.09
RF positivity	1.72 (0.82 – 3.58)	0.15	2.23 (0.97-5.12)	0.06
aCCP positivity	1.62(0.84-3.13)	0.15		
RF and aCCP positivity	1.65(0.87-3.14)	0.13		
Age per year	1.01(0.99-1.04)	0.28		
Monitor strategy A B C D	A. 1.00 (-) B. 0.53 (0.15-1.85) C. 0.96(0.31-3) D. 1.29 (0.43-3.22)	A. NA B. 0.27 C. 0.95 D. 0.63	A. 1.00 (-) B. 0.24 (0.06-0.98) C. 0.66 (0.19 – 2.22) D. 0.96 (0.31-2.98)	A.NA B. 0.05 C. 0.5 D. 0.94
Symptom duration (mo.)	1.0 (0.94 – 1.1)	0.85		
SvH units	1.04 (0.93 – 1.16)	0.47		
Smoking	1.56 (0.73 – 3.4)	0.25		
Alternative HAQ	1.16 (0.71-1.9)	0.55		
USJC per point	0.98 (0.92 – 1.05)	0.68	0.96 (0.89 – 1.04)	0.31

A

ADDENDUM
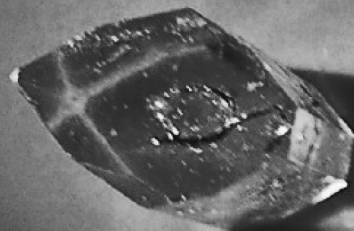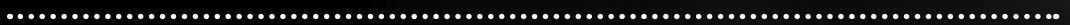
PhD PORTFOLIO

Name: David Folkert ten Cate
Erasmus MC department: Rheumatology
Research school: NIHES
PhD period: September 2009 – April 2015
Promotor: Prof. Dr. J.M.W. Hazes
Co-promotors: Dr. J.J. Luime
                           Dr. J.W.G. Jacobs

General academic and research skills

2010
BROK (GCP) Erasmus MC, Rotterdam

2014
Refresher course BROK (GCP) Erasmus MC, Rotterdam

In depth courses

2009-2010
Three ultrasound 'hands on' courses, Houten

2010-2013
Master of Science, specialization Clinical Epidemiology, NIHES, Erasmus MC, Rotterdam

(Inter)national conferences

2009
- REUS-meeting (presently IRON) (Imaging working group of Nederlandse Vereniging voor Reumatologie (NVR)), Eindhoven (oral presentation)
- STIZORO-meeting (presently Cicero) (collaboration of rheumatologists in the Southwest of the Netherlands), Rotterdam (oral presentation)

A

2011
- Nederlandse Vereniging voor Reumatologie (NVR) Najaarsdagen, Papendal *(oral presentation)*
- Nederlandse Vereniging voor Reumatologie (NVR) Najaarsdagen, Papendal (oral presentation)
- Nederlandse Vereniging voor Reumatologie (NVR) Najaarsdagen, Papendal *(poster presentation)*
- Meeting of the American College of Rheumatology (ACR) Chicago *(poster presentation)*
- Meeting of the American College of Rheumatology (ACR) Chicago *(poster presentation)*

2009
- IRON meeting, Almere (oral presentation)

2012
- Nederlandse Vereniging voor Reumatologie (NVR) Najaarsdagen, Papendal *(poster presentation)*
- Meeting of the American College of Rheumatology (ACR) Washington (poster presentation)

2013
- Meeting of the European League Against Rheumatism (EULAR) Madrid *(poster presentation in poster tour)*
- Meeting of the American College of Rheumatology (ACR) San Diego *(poster presentation)*
- Nederlandse Vereniging voor Reumatologie (NVR) Najaarsdagen, Papendal *(oral presentation)*
- Nederlandse Vereniging voor Reumatologie (NVR) Najaarsdagen, Papendal *(poster presentation)*

Teaching activities

2009
- Supervising research internship of 4th year medical student

2010
- Bedside teaching for general practitioners, Erasmus MC

- Supervising ultrasound rotation of an intern, Erasmus MC
- Supervising two 1st year medical students how to write a systematic review
- Supervising research internship of 4th year technical student

2011
- Supervising ultrasound rotation of three interns, Erasmus MC
- Supervising research internship of 4th year technical student

2012
- Supervising ultrasound rotation of an intern, Erasmus MC
- Supervising research internship of 4th year medical student
- Supervising research internship of 5th year medical student
- Supervising two 1st year medical students how to write a systematic review

2013
- Workshop 'Musculoskeletal Ultrasonography': 2nd year medical students, Erasmus MC
- Plenary session 'Musculoskeletal Ultrasonography': 3rd year medical students, Erasmus MC

Other

2009-2013
- Organising semiannual calibration sessions for ultrasonographers of Echografie in Vroege Artritis (EVA) working group

2010
- Registered ultrasonographer, certified by the imaging working group 'REUS' (presently IRON) of the Nederlandse Vereniging voor Reumatologie (NVR)

2010-2013
- Outpatient clinic: Muskuloskeletal ultrasound, Rheumatology, Erasmus MC

2012
- Study tour 'How to take synovial biopsies using ultrasound', Policlinico San Matteo, Pavia, Italy
- Advisory council, Targeted ultrasound initiative, Utrecht

A

2012-2013
- Management of research nurses

PUBLICATIONS

Publications (this thesis)

Ten Cate DF, Luime JJ, Hazes JM, Jacobs JW, Landewe R. Does the intraclass correlation coefficient always reliably express reliability? Comment on the article by Cheung et al. Arthritis Care Res (Hoboken). 2010;62:1357-8; author reply 8.

Dopazo Gonzalez N, **Ten Cate DF**, Swen WA, Mera Varela A, Insua Vilarino SA, Perez-Pampin E, et al. The most reliable probe position in the ultrasonographic examination of the wrist in rheumatoid arthritis. Clin Exp Rheumatol. 2012;30:566-9.

Ten Cate DF, Luime JJ, van der Ven M, Hazes JM, Kooiman K, de Jong N, et al. Very different performance of the power Doppler modalities of several ultrasound machines ascertained by a microvessel flow phantom. Arthritis Res Ther. 2013;15:R162.

Ten Cate DF, Luime JJ, Swen N, Gerards AH, De Jager MH, Basoski NM, et al. Role of ultrasonography in diagnosing early rheumatoid arthritis and remission of rheumatoid arthritis--a systematic review of the literature. Arthritis Res Ther. 2013;15:R4.

Ten Cate DF, Luime JJ, Hazes JM, Kleinrensink GJ, Jacobs JW. Is the frequent sonographic anechoic area distally in metacarpophalangeal joints a sign of arthritis? Ultrasound Med Biol. 2014;40:2537-41.

Ten Cate DF, Glaser N, Luime JJ, Lam KH, Jacobs JW, Selles R, et al. A comparison between ultrasonographic, surgical and histological assessment of tenosynovits in a cohort of idiopathic carpal tunnel syndrome patients. Clin Rheumatol. 2014.

Jacobs JW, **Ten Cate DF**, van Laar JM. Monitoring of rheumatoid arthritis disease activity in individual patients: still a hurdle when implementing the treat-to-target principle in daily clinical practice. Rheumatology (Oxford). 2014.

Ten Cate DF, Jacobs JWG, Swen WAA, Hazes JMW, Jager MH de., Basoski NM, Haagsma CJ, Luime JJ*, Gerards AH* * both authors contributed equally. Additional value of baseline ultrasonography in early RA to predict failure to reach DAS 28 remission after 1 year of tight-control treatment. Submitted

Onna van M., **Ten Cate DF**, K. L. Tsoi MD, A. J. L. Meier, J. W. G. Jacobs, Westgeest AAA, Meijer PBL, Beek van MC, Rensen WHJ, Bijlsma JWJ. Assessment of disease activity in patients with rheumatoid arthritis using OST measurements, a non-invasive operator-independent imaging technique. Submitted

Publications (other)

Ten Cate DF. 'Does the TRIPs Agreement deserve the blame for limited access to drugs and medical technology?'. Medicus Tropicus. 2007;45:14-6.

A

Takkenberg RB, Zaaijer HL, **ten Cate DF**, Biemond BJ, Jansen PL, Reesink HW. [Reactivation of hepatitis B virus in patients with hematological neoplasms] Reactivatie van het hepatitis B-virus bij patienten met een hematologische maligniteit. Ned Tijdschr Geneeskd. 2009;153:926-31.

Ten Cate DF, Gerards A, Haagsma C, Hazes JMW, Jacobs JWG, Swen WAA. 'Startschot klinkt voor nieuw echoproject'. NTvR. 2009;4:75-6.

ABOUT THE AUTHOR

In zijn jonge jaren was het gissen wat David ten Cate (geboren 4 september 1981 te Rotterdam) later zou gaan doen met zijn leven. Deze veelzijdige jongeman was als kind al in de meest uiteenlopende onderwerpen geïnteresseerd. Van sterren kijken tot archeologie en van de oudheid tot aan de bijbel, David wist over elk onderwerp wel wat leerzaams te vertellen. Gelukkig hield hij ook nog tijd over om te hockeyen, judoën en gitaar te spelen, het een met wat meer succes dan het ander.

Na zijn middelbare school op het Erasmiaans Gymnasium te Rotterdam, bleek het kiezen van een studie voor David een lastige opgave. Een kort uitstapje aan de Universiteit Leiden (Faculteit Sterrenkunde) in 1999, was niet wat hij ervan verwachtte, en het volgende jaar deed David mee aan de loting voor de studie geneeskunde. Helaas zou hij 3 maal worden uitgeloot, alvorens via de decentrale selectie in 2003 dan eindelijk met de studie te kunnen beginnen in het AMC.

In de tussenliggende jaren stortte David zich op zijn andere grote liefdes, de geschiedenis van de oudheid en de bijbel. Wederom met een vertraging sloot hij dit traject af met een bachelor in de geschiedenis aan de Universiteit Utrecht in 2006.

In Amsterdam bleek tijdens de geneeskundestudie opnieuw dat keuzes maken niet Davids sterkste kant is. Na een kort avontuur bij Religiestudies en een wetenschaps-stage bij de maag-darm-lever (MDL) ziekten, verloor hij uiteindelijk zijn hart aan de reumatologie, waarin hij eerst wilde promoveren. Hiermee werd in 2009 begonnen en resulteerde in dit proefschrift, begin 2015. In de tussentijd startte hij op 1 september 2013 met de vooropleiding interne geneeskunde in het Ikazia ziekenhuis te Rotterdam, onder opleider Dr. A.A.M. Zandbergen. In 2017 begint hij de formele opleiding tot reumatoloog onder Prof. Dr. J.M.W. Hazes.

Zijn moeder is bijzonder trots.

David woont samen in Rotterdam met Annemieke. Samen hebben zij twee prachtige jongens, Luuk en Gijs.

A

DANKWOORD

Het schrijven van een proefschrift is zo'n enorm project; afhankelijk van zoveel factoren en van zoveel mensen, dat het ondoenlijk is om eenieder de eer te geven die hem of haar toekomt. Dit maakt het meest gelezen deel van een proefschrift direct het minst complete deel.

Sta mij toe te beginnen alle patiënten te bedanken die geheel belangeloos tijd en energie hebben gestoken in mijn proefschrift. Ondanks dat de medische wetenschap zich de afgelopen jaren niet altijd van zijn beste kant heeft laten zien kwam het maar zelden voor dat patiënten weigerden mee te werken. Dank u zeer, zonder u is de medische wetenschap nergens.

Direct hierna wil ik mijn promotor, professor Hazes, bedanken. Ik vind het erg bijzonder dat u altijd tijd had voor overleg, ook als die tijd er eigenlijk niet was. Nooit heb ik tijdens gesprekken het idee gehad dat 'we moesten afronden'; nooit heb ik u met een schuin oog naar de klok zien kijken, hoewel uw agenda altijd overboekt is. U gaf mijn onderzoek op de een of andere manier altijd op het juiste moment de juiste richting. Verder, het vertrouwen dat u mij heeft gegeven om de echografie in de kliniek op te zetten, terwijl ik er in het begin ook niks van wist, maakt bijna verlegen. Het klimaat dat u aldus op uw afdeling gecreëerd heeft geeft enorm veel rust voor jonge onderzoekers om te groeien en op momenten boven zichzelf uit te stijgen. U hebt mij het vak 'wetenschap' geleerd; ik kijk er erg naar uit om ook het vak 'reumatologie' van u te leren.

Mijn co-promotor, Jolanda Luime. Jolanda, de eerste keer dat ik met je sprak, op je kamer, stond daar de eerste echomachine van de afdeling. We keken ernaar (hoe gaat ie eigenlijk aan?), en besloten dat we met elkaar dit spannende project zouden aangaan. Een project dat we van de grond af moesten opbouwen. Ik denk dat het een groot succes is geworden, met verschillende nieuwe echo-onderzoeken buiten mijn promotieonderzoek, die al tijdens mijn promotietraject zijn gestart. In de dagelijkse begeleiding heb je me behoed voor te snel en dus te onzorgvuldig handelen. Ook je statistische kwinkslagen hebben mijn proefschrift beter gemaakt.

Ik ben in de gelukkige omstandigheid dat ik nóg een co-promotor heb, en wat voor een! Hans Jacobs. Hans, het is mij nog steeds een raadsel hoe jij manuscripten zo snel als een afwezigheidsmelding wist terug te sturen en dan ook nog eens met pagina's extra aan commentaar. Je hebt een afwezigheidsmelding trouwens, nu inclusief de taal van 'het oude volk'. Mooi. Je hebt mij veel geleerd, maar ik zal me vooral je inspanningen om mij een artikel te leren schrijven herinneren. Het belang daarvan kan niet overschat worden. Ik waardeer het erg hoeveel tijd en energie je in mij hebt geïnvesteerd.

A

All co-authors, in alphabetical order. Natalja Basoski, Michiel van Beek, Marietta Bertleff, professor Bijlsma, Hans Bosch, Noelia Dopazo Gonzalez, Andreas Gerards, Nick Glaser, professor Gómez-Reino, Cees Haagsma, professor Hazes, Dr. Insua Vilariño, Hans Jacobs, Mike de Jager, Nico de Jong, professor Kleinrensink (unfortunately, no Nobel prize), Klazina Kooiman, professor van Laar, Dr. Lam, professor Landewé, Jolanda Luime, Louise Meier, Dr. Meijer, Dr. Mera Varela, Marieke van Onna, Dr. Perez-Pampin, Wouter Rensen, Ruud Selles, Nanno Swen, Ka Lai Tsoi, Myrthe van der Ven, Dr. Westgeest.

De leden van de 'Echografie bij Vroege Artritis' (EVA)-club, Hans Jacobs, Nanno Swen, Jolanda Luime, professor Hazes, Natalja Basoski, Andreas Gerards, Cees Haagsma en Mike de Jager; enorm bedankt voor de praktische medewerking aan het onderzoek. Voor de echografisten nog in het bijzonder dat jullie tijd hebben gevonden om ten behoeve van mijn onderzoek tussen jullie razend drukke poli's door zoveel echo's te verrichten. Een prestatie van formaat. Dank.

De kleine commissie, professor de Jong, professor Kleinrensink en professor Landewé. Dank dat u mijn proefschrift wilde beoordelen.

Verder, professor Bloem, dank dat u in de grote commissie wilde plaatsnemen. Als hoogleraar in de radiologie kunt u als geen ander dit proefschrift in een groter kader plaatsen.

Dr. Wakefield, it is a great honour for me that you as an internationally known sonography expert agreed to participate in the thesis committee. I hope you enjoyed reading the thesis, my defence of it and your time in the Netherlands. I look forward to meeting you again at conferences.

Als laatste lid van de grote commissie, Nanno Swen. Nanno, je onbegrensde enthousiasme voor de echografie is aanstekelijk. In een ver verleden, tijdens mijn coschap kindergeneeskunde in Alkmaar, bedacht ik dat ik reumatoloog wilde worden. Toen ik eens met een willekeurige reumatoloog, ene Dr. Swen, uit dat ziekenhuis ging praten hadden we beiden niet kunnen bevroeden dat jij 8 jaar later in mijn promotie-commissie zou zitten.

Om de EVA studie een succes te maken moest de logistiek op orde zijn. Dit kan niet zonder een team van logistieke geniën (researchnurses/polimedewerkers/datamanagers etcetera) die altijd bereid waren tot een beetje schuiven in de agenda. In onderzoek gaan de dingen nooit zo als ze moeten; het moet altijd maar zoals het gaat, en dankzij jullie kon dat. Dank (ik realiseer me dat deze lijst bij uitstek incompleet is), Sjaan, Conny, Anke, Anneke, Hetty, Ron, Naomi, Trudy, Tania, Gerda, Mireille, Annemieke, Esther, Inge en Coby.

Joyce de Jong, als moeder gans van de afdeling, bedankt voor je steun en meedenken; ook fijn dat er altijd snoep was, dan gaat het denken meteen een stuk beter.

Mijn persoonlijke leescommissie. Marijn; mooie tijden in Pavia. Nooit cappuccino bij het avondeten. Martijn, dank dat je me altijd weer met een of andere statistische truc uit de brand hielp. Je proefschrift wordt vast briljant en onbegrijpelijk voor de mindere STATA-goden. Myrthe, wat leuk dat het echografisch onderzoek in Rotterdam een vervolg heeft gekregen. Een heel mooi vervolg met binnenkort jouw promotie; ik blijf je na mijn promotie op de voet volgen.

Mijn collega-promovendi/kamergenoten. De afdeling reumatologie is de afgelopen jaren hard gegroeid. Dus we zijn talrijk! Annelieke, Celien, Esther, Florentien, Hilal, Jenny, Lonneke, Maren, Marie Louise, Martijn, Myrthe, Pascal, Sjel. Wat een mooie afdeling hebben we. Dank voor alle ondersteuning bij de voltooiing van mijn promotie en vooral voor de gezelligheid. Soms werd het een beetje te gezellig en werd het gekakel in het kippenhok me wat teveel. Achteraf weet ik, gekakel om je heen is een voorwaarde om een proefschrift af te ronden. Verder van de afdeling reumatologie, Radboud, sinds ik jou ken let ik altijd op 'confounding'. Philomine, nog even en jij bent ook gepromoveerd. Gaan we dan echo's maken op de polikliniek kinderreumatologie?

Dr. Zandbergen, beste Adrienne, wat een geluk dat ik jou trof als opleider in het Ikazia. Zonder jouw medewerking was dit proefschrift er niet gekomen. Assistenten van de interne geneeskunde in het Ikazia, dank voor de steun en de geweldige werksfeer. Topcollega's!

Jur, Kris, Alex, Marcus, Bou. Jaha, ik heb nu eindelijk eens een studie afgemaakt. De, altijd slechte, en altijd keiharde grappen in de app zijn tijdloos. Binnenkort weer tijd voor wat meer van dat soort grappen in persoon.

Baden, lunch?

Els, wat breng jij een rust. Blijf je nog lang bij ons?

Wim en Elly, wat zijn jullie een leuke opa en oma voor onze jongens. Bedankt dat jullie altijd klaar stonden om in te springen als ik weer eens aan het werk was.

Lieve papa en mama. Dé levensles die ik van jullie geleerd heb is: "Kan niet bestaat niet." Deze mantra heeft mij gevormd. Hierdoor ben ik misschien wel wat eigenwijzer geworden dan de bedoeling was, maar dank dat ik altijd op jullie kon (en kán!) terugvallen, ook als mijn keuzes wat minder gelukkig uitpakten. Bij jullie is het, en blijft het, écht thuiskomen. Ik geef de mantra door aan jullie lieve kleinkindjes en hoop dat zij net zoveel steun, liefde en vertrouwen van mij ervaren als dat ik van jullie heb ervaren.

Pommetje, mijn lieve zusje. Het is iedere keer weer verfrissend om jouw blik op de zaak, welke zaak dan ook, te krijgen. Samengevat, pfff, boeie. Had ik maar wat meer van jouw instelling.

Mijn, paranimfen. Allereerst Tim, je bent echt 'mijn grote broer'. Je adviezen de afgelopen jaren waren altijd raak. Dank dat je mij terzijde wilt staan tijdens de verdediging. Boudewijn, historisch is het een van de taken van de paranimfen om ervoor te zorgen dat de promovendus niet door het publiek wordt gelyncht als hij hen onwelgevallige

A

dingen zegt. Wat dat betreft heb ik het volste vertrouwen in je. Dat je altijd, zelfs vanuit Somalië, interesse toonde in dit jarenlange project vind ik heel bijzonder.

Mijn vriendinnetje, mijn lieve mooie Miekie. De buitenwereld weet niet half wat je met me te stellen hebt gehad de afgelopen jaren. En dat laten we maar zo. Steeds vaker vroeg ik me af: "Hoeveel mag werk kosten?". Steeds vaker werd het antwoord: "Niet zo veel." Dat jij die vraag nooit hardop gesteld hebt, ook niet toen de situatie schreeuwde om een antwoord, is de enige reden dat dit proefschrift is voltooid. Ik ga je daar in woorden nooit genoeg voor kunnen bedanken. Liefie, het is volbracht, eindelijk hebben we weer meer tijd voor elkaar; jij bent het, die mij ik maakt.

Luuk en Gijs, op het moment dat ik dit schrijf begrijpen jullie nog niets van wat er gaande is. (Is papa nu weer werken? Waardoor is papa soms zo gestresst?) Jullie onvoorwaardelijke liefde maakt nietig. De donderdagen zijn eindelijk weer helemaal van ons, jongens. Tegen de tijd dat jullie dit zelf kunnen lezen, mijn promotie een verre echo uit het verleden, weet dan, dat jullie bij me waren, met elke toets die ik op het toetsenbord aansloeg. Zingen we ons liedje, jongens? 'Hé, hallo, lieve kleine Luuk en Gijs, jullie zijn mijn mooie zoontjes, ik ben zo gek op jou en jou. Je bent zo mooi en jij bent dat ook, ik hou zoveel van jullie.'