Answer questions and earn CME https://education.aanem.org/URL/JR82.

INVITED REVIEW

Muscle ultrasound: Present state and future opportunities

Juerd Wijntjes MD | Nens van Alfen MD, PhD

Department of Neurology and Clinical Neurophysiology, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands

Correspondence
Juerd Wijntjes, Department of Neurology and Clinical Neurophysiology, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands.
Email: juerd.wijntjes@radboudumc.nl

Abstract
Muscle ultrasound is a valuable addition to the neuromuscular toolkit in both the clinic and research settings, with proven value and reliability. However, it is currently not fulfilling its full potential in the diagnostic care of patients with neuromuscular disease. This review highlights the possibilities and pitfalls of muscle ultrasound as a diagnostic tool and biomarker, and discusses challenges to its widespread implementation. We expect that limitations in visual image interpretation, posed by user inexperience, could be overcome with simpler scoring systems and the help of deep-learning algorithms. In addition, more information should be collected on the relation between specific neuromuscular disorders, disease stages, and expected ultrasound abnormalities, as this will enhance specificity of the technique and enable the use of muscle ultrasound as a biomarker. Quantified muscle ultrasound gives the most sensitive results but is hampered by the need for device-specific reference values. Efforts in creating dedicated muscle ultrasound systems and artificial intelligence to help with image interpretation are expected to improve usability. Finally, the standard inclusion of muscle and nerve ultrasound in neuromuscular teaching curricula and guidelines will facilitate further implementation in practice. Our hope is that this review will help in unleashing muscle ultrasound’s full potential.

KEYWORDS
biomarker, diagnostic screening, implementation, muscle ultrasound, neuromuscular ultrasound

1 | INTRODUCTION

Muscle ultrasound is a valid screening tool for neuromuscular disease.1–3 With ultrasound we are able to detect pathological changes in neuromuscular disease that reflect the fatty replacement and fibrosis of affected muscles.4 It is a patient-friendly and non-invasive technique, that can easily be used in an outpatient setting and at the bedside. Muscle ultrasound has been proven to be a reliable alternative to more invasive investigations such as EMG in screening for neuromuscular diseases in children.5 Most commonly, muscle ultrasound images are evaluated visually or semi-quantitatively using the so-called Heckmatt scale.6 In addition, some centers quantify the image results by analyzing mean grayscale level or backscatter values. The reported sensitivity for these techniques to detect a neuromuscular disorder varies from 70% for pure visual analysis up to 92% for a fully quantified approach.2,6,7 Muscle ultrasound performs less well as a screening tool in metabolic myopathies or in children under 3 years of age, that is when there are no or just a limited amount of structural muscle abnormalities.8,9

While visual or quantitatively evaluated echogenicity is the most commonly reported clinical muscle ultrasound parameter, ultrasound
Muscle atrophy is easily assessed using the caliper function on ultrasound machines to measure muscle thickness. Doppler blood flow signal can show muscle vascularization, and finding an increased intramuscular blood flow may be useful to assess inflammatory activity in patients with myositis. Changes in muscle architecture caused by fibrosis and fatty degeneration in neuromuscular disease do not only lead to increased grayscale levels, but also affect tissue elasticity and muscle anisotropy. To depict changes in the mechanical properties of muscle, ultrasound techniques such as strain elastography, shear wave elastography and viscoelastic response imaging have been developed. A progressive increase in muscle stiffness was shown in boys with Duchenne muscular dystrophy, making elastography a potential biomarker for research on the treatment response in muscular dystrophy. Extracting higher order features from the muscle ultrasound image, such as local tissue anisotropy, can also help identify abnormal muscle tissue architecture in a way that might be less dependent on the ultrasound machine settings.

Ultrasound can also capture muscle movements, such as fasciculations and fibrillations. This dynamic use of muscle ultrasound provides enhanced sensitivity for the detection of fasciculations, and can improve diagnostic certainty in patients with ALS. To facilitate the use of muscle ultrasound videos in the clinic, the frame by frame analysis of movements can be made more efficient by using a computer algorithm. A more advanced option to quantify tissue motion is speckle tracking, a common technique in cardiac imaging, that is currently used to improve the diagnosis of diaphragm pathology. With the advent of new treatment options for neuromuscular disorders, such as nusinersin for spinal muscle atrophy, there is a growing need for sensitive, relevant and accessible biomarkers that can be used in future treatment trials. Muscle ultrasound could serve as such a research biomarker, as it is non-invasive, portable and patient-friendly. Like muscle MRI, it shows muscle tissue architecture and pathological tissue changes that may substitute for histological measures and obviate the need for repeated tissue biopsy during study follow up. Several studies report promising results for the use of muscle ultrasound as such a biomarker in muscular dystrophies, showing high correlations with disease progression and functional performance over time.

The addition of muscle ultrasound to the neuromuscular medicine toolkit and the electrodagnostic lab has the potential to replace the existing diagnostic modalities in specific situations. Yet, promising as it is, muscle ultrasound imaging is currently not fulfilling its full potential in the diagnostic care of patients with neuromuscular disorders. The two most important drawbacks currently appear to be that on the one hand qualitative muscle ultrasound analysis is very dependent on observer experience, and that quantitative analysis needs reference values that are highly device-dependent and require a lot of time and effort to acquire. Thus, muscle ultrasound is currently only routinely used by specialized clinicians in centers with a large neuromuscular case load that makes gathering sufficient experience in visual grading and investing in collecting reference values worthwhile. Additionally, we also need more evidence from studies to show how muscle ultrasound alters clinical decision making, improves patient outcome and is cost-effective, either as a stand-alone diagnostic tool or as an integrated part of a stepwise diagnostic testing strategy that starts with ultrasound as the most patient-friendly test.

In our opinion, the lack of widespread implementation of this technique among clinicians who diagnose and treat neuromuscular patients is unfortunate. A targeted effort will be needed to remove the obstacles that hamper muscle ultrasound’s implementation and use, and to obtain regulatory approval for the use of muscle ultrasound as a biomarker in treatment trials. In this review, we want to highlight the current possibilities of muscle ultrasound as a diagnostic tool and biomarker for different neuromuscular disorders, and discuss the challenges that need to be overcome before the technique can be widely implemented and realize its full potential.

2 | CURRENT POSSIBILITIES OF MUSCLE ULTRASOUND

2.1 | Muscle ultrasound: A technique primer

In diagnostic ultrasound applications, very high frequency (“ultra-”) sound waves are sent into the body and change direction when they encounter transitions between different tissue types that have different acoustic properties. Sound can either be reflected back to the probe surface to show up as a pixel on the ultrasound screen or be deflected away from the probe or even scattered across the tissue layers, depending on the angle of insonation and the size of the tissue structure encountered. The more energy the sound waves have when they return to the probe, the brighter the resulting pixel will be on the screen. The longer it takes for them to get back to the probe, the deeper the ultrasound machine will think they originated. Because only reflections show up on the screen, it is very important to maximize the amount of sound that is reflected. This is done by keeping the probe as perpendicular to the tissue as possible. In case of muscle tissue, this means scanning at a 90° angle to an underlying bone or large fascial structure, such as an interosseus membrane. Doing so will result in the typical appearance of muscle dubbed the “starry night appearance” in transverse scans, with relatively black looking muscle fibers (the black of the night) interspersed with white fascial structures (the stars). In longitudinal images, the muscle fiber direction and pennation angles are visualized.

The use of muscle ultrasound in neuromuscular disease is based on the principle that diseased muscles show changes in their texture, or tissue architecture, as healthy muscle fibers degenerate and are replaced by fibrosis and fatty infiltration. This mixing of different tissue types leads to an increase in the amount of ultrasound reflections, and therefore an increasingly brighter image appearance. Also, diseased muscles will often myotomal pattern of muscle, which is easily measured with muscle ultrasound.

As different neuromuscular disorders lead to different pathological changes in muscle histology, their appearance on muscle
**FIGURE 1** Properties of sound reflections (A). Transverse and longitudinal muscle ultrasound image of the gastrocnemius muscle of a healthy person (B). Ultrasound reflections in healthy (left) and diseased (right) muscle (C) and its histologic correlation (D). Figure 1D represents a microscopic image of a muscle biopsy, which is of a smaller scale than what is seen with ultrasound.

**FIGURE 2** Different pathological changes on muscle ultrasound in different neuromuscular disorders. Normal "starry night" appearance of a healthy biceps brachii muscle (A). Ground glass appearance of the tibialis anterior muscle in a patient with facioscapulohumeral dystrophy (FSH) (B). Attenuation of the ultrasound beam in the rectus femoris muscle of an FSHD patient (B,C). Focal areas of increased echogenicity (arrow) in the flexor carpi radialis muscle of a dermatomyositis patient (D). See-through appearance of the deltoid muscle of a patient with suspected myositis (E). Inflammation of the pretibial skin and subcutaneous layer (arrow) of a dermatomyositis patient (F). Patchy-appearing biceps brachii muscle with overall high echogenicity of a patient with long-standing myositis (G). Calcifications (arrowheads) with typical posterior acoustic shadowing in the rectus femoris muscle of a dermatomyositis patient (H). Thickening of fascia (arrow) in the rectus femoris muscle of a patient with eosinophilic fasciitis (I). Affected deep finger flexor muscles (arrow) in a patient with inclusion body myositis (J). Unaffected and affected forearm muscles in a young girl with left arm peripheral nerve trauma (K). Moth-eaten pattern (arrowheads) in the interosseous dorsalis muscle of a patient with long-standing neurogenic disease (L).
ultrasound will also differ. This leads to several patterns of change in muscle ultrasound images that can help distinguish these disorders from one another.

Muscular dystrophies are associated with extensive replacement of muscle fibers by fat and fibrosis as the disease progresses. Typically, this will result in a homogeneously increased grayscale level of the whole muscle with loss of architectural features, that has been dubbed a “ground glass” appearance (Figure 2B). When pathology progresses, the top layers of an affected muscle will reflect much of the ultrasound. This so-called attenuation of the ultrasound beam will result in a very bright top muscle layer with a dark area underneath (Figure 2C). Depending on the type of dystrophy and the specific muscle imaged, atrophy may or may not be present in dystrophic muscles. Typically, no atrophy is found in the younger Duchenne muscular dystrophy (DMD) patients, and in facioscapulohumeral dystrophy (FSHD) atrophy is commonly found only in selected muscles such as the rectus femoris or trapezius, but not in many others. Also depending on the type of dystrophy, the muscle changes may be found throughout the whole muscle, or in specific regions only that expand over time (such as in FSHD). Other types of neuromuscular disorders such as late-onset Pompe disease could likewise affect select muscles or affect only parts of muscles.

Inflammatory myopathies also show specific changes on muscle ultrasound images. In dermatomyositis and polymyositis, the abnormalities usual start out as focal areas of increased echogenicity (Figure 2D), that generalize with progression of the disease. In an acute myositis flare, muscle edema can be seen as an overall echogenicity increase without attenuation, that has been dubbed a “see-through echogenicity increase” (Figure 2E). However, muscle ultrasound is (much) less sensitive to edema than for example MRI with short tau inversion recovery sequences. Inflammation of the skin and subcutaneous layer may be found in dermatomyositis and in cases with myositis in diffuse connective tissue disease (Figure 2F). In long-standing or chronic myositis, atrophic muscles with high echogenicity levels are common, in which the image abnormalities are usually still somewhat patchy, as some areas with intact muscle fibers remain (Figure 2G). In juvenile dermatomyositis patients, subcutaneous or intramuscular calcifications can be easily detected with ultrasound (Figure 2H). In (eosinophilic) fascitis, irregular fascial thickening with fuzzy fascial edges is seen (Figure 2I). Inclusion body myositis (IBM) is a distinct sub form of inflammatory myopathy that occurs in older patients and often prominently involves the quadriceps and deep finger flexor muscles (Figure 2J). The distinct grayscale contrast of the flexor digitorum profundus and adjacent flexor carpi ulnaris muscle was found to be a sensitive diagnostic indicator of IBM.

Neurogenic disorders have a variable appearance on muscle ultrasound that will depend on the severity of the axonal loss, the duration and progression of the disorder, and on whether reinnervation has occurred. In our experience, three main patterns can be discerned. Usually no abnormalities are found in cases of slight monophasic axonal injury, such as a mild radiculopathy, where collateral reinnervation can occur within 3-4 months. As such, muscle ultrasound does not seem to be a good screening instrument for these types of pathology. In cases with longer standing denervation and none to some, but still incomplete reinnervation, such as in Sunderland grade III-V nerve trauma, muscle ultrasound will show a patchy to diffuse increase in echogenicity. This type of abnormality is usually easy to pick up visually when comparing the affected and unaffected limb (Figure 2K). It can help pre-screen muscles of interest for needle EMG, for example in children or other patients who do not tolerate needling well. The myotomal pattern of muscle involvement as determined by muscle ultrasound can be used to assess mononeuropathies and radiculopathies. The third pattern of muscle ultrasound changes in neurogenic disorders is the most characteristic. It has been dubbed the “moth-eaten” pattern (Figure 2L), and it consists of round, dark areas that contain remaining viable, usually enlarged motor units, surrounded by tissue with increased echogenicity that reflects permanent denervation and fibrosis. This pattern is seen in long-standing and progressive neurogenic diseases such as spinal muscular atrophy or persistent radicular compression. In patients with ALS, peripheral motor neuron involvement can show up as any of the patterns above, depending on the disease stage and extent of denervation of that specific muscle. At the initial presentation, the most common findings are some muscles with an echogenicity increase and a few muscles with atrophy. More atrophic and strongly hyperechogenic muscles will be found as the disease progresses. Most conspicuous though at onset are the fasciculations that occur in about 50% of the patients at the time of presentation and can be very well observed with ultrasound.

2.2 | Using muscle ultrasound: Visual evaluation

The easiest way to use muscle ultrasound is by scanning a muscle and looking at the image or video, either at the bedside or offline. Simple visual analysis provides a lot of information about the overall muscular echogenicity, texture, and anatomical context. Taking the subcutaneous fat layer as a reference, clearly abnormal muscles are easy to spot (ie Figure 2B). But sometimes changes are not that obvious or more difficult to interpret. Most of the 200+ skeletal muscles in our body have different architectures and different ratios of muscle to connective tissue content.

Muscle tissue content, and hence echogenicity, will also change significantly with changes in several physical parameters, such as weight, age, hand-dominance (for upper extremity muscles) and nutritional status. Muscle echogenicity will not change during childhood and adolescence, even as muscle diameter increases, but in adults over 50 years the grayscale levels will progressively increase because of sarcopenia, that is the age-related increase in fibrotic content of muscle. Obesity leads to an increase of both subcutaneous and intramuscular fat content. Even with attenuation of the ultrasound beam through a thicker subcutaneous fat layer, the net effect for the leg and abdominal muscles will be in increase in echogenicity. Adult males usually have larger muscle bulk and muscle fibers, and the latter will slightly lower overall echogenicity as there are less tissue transitions per surface area.
Interpreting the visual evaluation of muscle texture and grayscale levels thus strongly depends on the subject and observer experience. In practice, this limits the sensitivity for making a visual distinction between normal and diseased muscle to around 70%.\textsuperscript{2,7}

To make visual evaluation of muscle ultrasound more objective, Heckmatt and co-workers developed a four-point visual grading scale, that classifies images based on the muscle grayscale level compared to the overlying subcutaneous fat layer, the presence or absence of a distinct muscle architecture, and the amount of attenuation leading to decreased visibility of the underlying bone or fascia echo (Figure 3).\textsuperscript{6} The reported sensitivity of this technique for detecting neuromuscular disease ranges from 71% to 76%.\textsuperscript{2,44,45}

\section*{2.3 | Using muscle ultrasound: Quantitative evaluation}

Currently, the most sensitive and validated approach for using muscle ultrasound to discriminate between healthy and diseased muscle is offered by quantitative muscle ultrasound (QMUS) echogenicity analysis, using either the mean echogenicity or a calibrated backscatter technique.\textsuperscript{46-48} The first technique calculates the overall mean grayscale level within a manually selected region of interest in the ultrasound image, and compares this to a reference value for that specific muscle that is corrected for the influence of age, length, sex and weight.\textsuperscript{40} The size of this region of interest matters, as it can be shown theoretically that the larger the ROI size (within the muscle boundaries), the less variable the grayscale level results will be. Calculating the mean grayscale level can be done using the histogram function of freely available image software, such as ImageJ (https://imagej.nih.gov/ij/) (Figure 4). Using the reference values, the measured grayscale level is transformed into a Z-score, which is the number of standard deviations from the mean echogenicity for that muscle.\textsuperscript{7}

\[ Z = \frac{\text{measured value} - \text{normal value}}{\text{standard deviation of normal value}} \]

In backscatter analysis the backscatter level represents the amount of signal reflected by tissue back to the ultrasound transducer in decibels. To minimize any device-dependent influences, the ultrasound signal is calibrated by using of a tissue-mimicking phantom, a fixed region of interest and a stepwise increase in overall gain that allows the user to establish the relationship between ultrasound grayscale levels and backscatter values, provided that the ultrasound system manufacturer is willing to share proprietary ultrasound device data that links the gain increases to known changes in decibel levels. Calibrated backscatter
measurements can reliably quantify ultrasound signals of skeletal muscle and are more reproducible between different ultrasound systems than average grayscale levels because of the calibration. However, the need for proprietary equipment information can seriously limit the practical use of this technique.48-50

QMUS is currently the most reliable and sensitive technique for detecting neuromuscular disease, with a sensitivity of 92%.1 From our clinical experience, it will detect any patient with a classic neuromuscular disorder such as DMD, Pompe disease or spinal muscular atrophy. As these disorders are often already diagnosed by their phenotype and genetic testing, ultrasound is probably more suited here as a biomarker to monitor disease severity and progression. QMUS alone had a sensitivity of 96% and specificity of 84% for discriminating ALS from mimics.28 An overview of the diagnostic values for different indications can be found in Table 1.

2.4 | Using muscle ultrasound: Higher order muscle texture features

The development of more advanced techniques that analyze muscle texture is considered promising, because these so-called “higher order features” are not based on grayscale levels and might be less device-dependent.51 The analyses of higher order texture features is based on the use of spatial variation of pixel intensities, that reflects the microstructure of the muscle imaged. Texture analysis uses the mathematical calculation of relative relationships of adjacent pixels. It can for example measure the number of times in which one pixel of a given grayscale level is found adjacent to another pixel with a different grayscale level to extract various texture features from this, such as contrast, entropy and correlation.52 Preliminary work showed that texture analysis can reliably distinguish between neurogenic and myogenic disease, and can for example distinguish ALS-affected muscles from those of healthy controls.53,54

Another recently developed technique made use of the degree of uniformity in muscle fiber orientation in the longitudinal direction, assessing texture anisotropy.52 Patients with DMD had lower texture anisotropy than healthy controls, and that the quantified texture anisotropy was much less affected by gain settings of the ultrasound machine than quantified grayscale levels.

2.5 | Dynamic muscle ultrasound

While static muscle ultrasound images are already very useful for diagnostic and follow up purposes in neuromuscular disease, ultrasound also brings a dynamic capability that can capture movement in a video sequence. For example, ultrasound videos can be used to show normal and abnormal muscle contraction patterns in muscular dystrophy, using a technique called speckle tracking.55 But probably the most useful dynamic application of muscle ultrasound is the detection of fasciculations in patients with suspected motor neuron disease. Because of the larger pick-up area of around 4 × 4-6 cm, that is the size of the muscle slice that is shown on the screen, muscle ultrasound has a higher sensitivity for picking up fasciculations than clinical observation or needle EMG.56 Fasciculations can be found in 10% to 30% of muscles that are subclinically involved or EMG-negative in patients with motor neuron disease.19,57 Combining static and dynamic ultrasound scanning for echogenicity changes and fasciculations with EMG provides 25% of the patients with amyotrophic lateral sclerosis with a more certain diagnosis at the time they first present.14,15 A small ultrasound screening protocol with ≥2/9 muscles positive for fasciculations, had a sensitivity and specificity of >90% for discriminating patients with clinically probable or definite ALS from disease mimics.26 Of note though, patients with other lower motor neuron disorders were excluded from this study cohort, limiting the clinical utility of this protocol.

The optimal scan time for detecting fasciculations has been found to be 30-60 seconds, similar to EMG.56,58,59 As a caveat, fasciculations of low amplitude on EMG (<1 mV) and a simple morphology with a maximum of four phases on EMG could be missed with ultrasound, as a recent study by Bokuda et al. showed. It is thought that these fasciculations occur in muscles that are just at the beginning of denervation.60,61

With the right equipment and settings, even fibrillations can be detected with ultrasound. A high-resolution linear ultrasound probe with a frequency range of 5-17 MHz will have a lateral resolution around 0.1-0.3 mm. As normal muscle fibers are around 0.04-0.06 mm in diameter, and atrophic denervated ones will likely be smaller, several fibers need to fibrillate and move simultaneously to be seen by the probe.62 Fibrillations last about 50–100 μs, so using a frame rate of 20 Hz it takes at least five fibrillations per second to cause enough displacement to be visible on screen. Just as for needle EMG, the occurrence of fibrillating muscle fibers is very temperature dependent, with a 30% decrease in their presence when the temperature is lowered from 39°C to 30°C.63,64

| Indication    | Author       | Age (years) | Number | Sensitivity | Specificity |
|---------------|--------------|-------------|--------|-------------|-------------|
| General screen| Pillen 2003  | 0.5-14      | 36     | 92%         | 90%         |
|               | Pillen 2006  | 0-18        | 76     | 87%         | 67%         |
|               | *Brandsma 2014 | 0-7    | 100     | 85%         | 75%         |
| IBM           | Noto 2014    | 68-79       | 18     | 100%        | NA          |
| ALS           | Arts 2012    | 23-79       | 59     | 96%         | 84%         |
| Metabolic     | Pillen 2006  | 0-15        | 53     | 25%-46%     | 85%-100%    |

Abbreviations: ALS, amyotrophic lateral sclerosis; IBM, inclusion body myositis.

*Visual or semi quantitative evaluation.

TABLE 1 Examples of diagnostic values of muscle ultrasound for screening and in specific neuromuscular disorders
3  |  IMPLEMENTING MUSCLE ULTRASOUND: CHALLENGES AND POSSIBLE SOLUTIONS

Even though the amount of scientific publications on muscle ultrasound has doubled over the last 10 years, the number of centers that routinely use the technique in clinical practice for general diagnostic and follow up purposes in neuromuscular disease is still quite limited. In this section we will discuss some of the technical, regulatory and setting-based issues that can underlie the underutilization of the technique. We will tentatively sketch steps that could be taken to realize a widespread implementation of neuromuscular ultrasound.

4  |  TECHNICAL CHALLENGES

The information captured in a muscle ultrasound image not only critically depends on system hardware, signal processing software settings and scanning technique, but also on our ability to process and correctly interpret all the information. Each of these variables comes with its own pitfalls, discussed below.

4.1  |  System hardware and ultrasound signal processing software settings

Medical ultrasound machines are sophisticated pieces of signal processing equipment. Ultrasound signals are sent out and captured
by arrays of piezo-electric elements in the ultrasound probes. These signals are amplified, and the soundwave energy is transformed from an analog into a digital signal that is then formed into a pixel image display on the screen. Just how exactly the image is put on the screen depends on multiple choices in hardware setup and software processing of the signals. Most of these choices are proprietary and made by the ultrasound manufacturer, but more can be added by the end user doing the scanning. This means that every ultrasound system will produce a different grayscale image of the same muscle (Figure 5A). As there is no uniform image standard for muscle ultrasound and the end user has only limited control over the image settings, this means that, currently, muscle images cannot directly be compared between different systems and settings.

For visual image analysis this implies that users will have to get familiar with their own systems and settings, and that comparing images from different systems will be hard, because of their inherent variability that can obscure a clear distinction between normal and pathologic tissue. Using a semi-quantitative grading system such as the Heckmatt scale may partially overcome these difficulties, but whether this is enough to use muscle ultrasound images from different sources in for example a multi-center diagnostic study or treatment trial has not been tested yet.

For quantitative image analysis that compares image grayscale or backscatter levels to a phantom or population-based reference value, the inherent image differences between devices and settings mean that new reference values are required for each type of device and setting used. Collecting these reference values is very labor-intensive. This currently precludes the use of QMUS for many centers, with the exception of those with large clinical or research caseloads and the means to collect these values.

Several strategies have been explored to circumvent the device-dependency of ultrasound reference values. Attempts have been made to find a fixed conversion equation between machines and settings, by scanning the same meat phantom with two different devices made to find a fixed conversion equation between machines and settings, this means that, currently, muscle images cannot directly be compared between different systems and settings.

In the clinical setting, ultrasound is mainly used to separate normal from abnormal muscle tissue, in order to identify patients with muscle disease or muscle involvement in another disorder (eg trauma, systemic disease). For clinicians with sufficient expertise, merely eyeballing the muscle image might be sufficient to already make this distinction (eg panel A vs B in Figure 4). Using the parameters of the Heckmatt grading scale helps to identify such abnormal muscle in ultrasound images. Reproducibility between observers or centers is of less concern. The Heckmatt grading scale scores are more suited for the assessment of the severity of muscle pathology or the follow up of patients over time. When the observer is uncertain if the ultrasound image is normal or not, or if small changes in grayscale levels need to be detected to use ultrasound as a biomarker for follow up or for detecting a treatment effect, results need to be reproducible between centers, quantification of the grayscale levels (or other ‘objective’ ultrasound parameters) becomes important.
Image interpretation using Heckmatt grading scale is very user friendly but comes with inherent pitfalls that limit its sensitivity. Visual evaluation of a muscle's grayscale level is sensitive to the so-called background effect. This optical illusion occurs when regions with identical grayscale levels become more or less conspicuous as the background levels change. This makes it harder to perceive the distinction from Heckmatt grade I (normal) to II (slightly abnormal), especially for muscles surrounded by other muscles that are affected differently, such as the flexor carpi radialis surrounded by the superficial and deep finger flexors (Figure 2J). Determining whether a muscle is Heckmatt grade I or II is also made more difficult by patient characteristics such as age (older people have increasingly brighter muscles) and body mass index (obese people have higher intramuscular fat content and therefore brighter muscles with less clear fiber texture) (Figure 5C,D).

Another disadvantage of the Heckmatt scale is that it was constructed in the early 1980s for the quadriceps muscle, with limited image quality compared to current ultrasound systems (Figure 5E). With the currently used (much) higher scan frequencies, more of the ultrasound beam will be scattered or attenuated by pathologic muscle tissue instead of transmitted or reflected as in the original classification. This leads to an earlier and more clearly discernible loss of tissue architecture, that will not always show up as an overall increase in echogenicity, because of the scattering effect (Figure 5F). The classification also assumes the presence of a large muscle with an underlying bony structure; features that are not present in many muscles that we can currently scan, such as the facial and submental muscles, abdominal muscles, calves, or the first dorsal interosseus muscle. Without an underlying bone or strongly reflective fascial structure, for example the interosseus membrane below tibialis anterior, the loss of a bone echo as one of the grading features of the Heckmatt scale that helps separate grade II, III and IV no longer applies. Small muscles such as the facial muscles may not be able to cause any significant beam attenuation through the diseased muscle tissue layer, precluding the use of another of the discerning Heckmatt grade features. Inherent variation of tissue architecture and hence grayscale levels in the 200+ different skeletal muscles in a human body also limit the standardized use of Heckmatt grading across all muscles.

All the factors above negatively influence visual muscle ultrasound assessment accuracy. As the distinction between Heckmatt grade I and grade II is what determines whether a muscle is normal or abnormal, translating to whether a neuromuscular disorder is present or not, the context-effect and patient characteristic pitfalls are the reason that the diagnostic screening sensitivity of visual assessment is stuck at about 75%. The other pitfalls mainly hamper a clear distinction between grade II-IV muscles, and can be problematic when one wants to use the Heckmatt scale for follow-up purposes. How difficult visual assessment and follow up will be, will most likely vary with the different types of neuromuscular pathology. Further, we have little information on the linearity of the test characteristics of the Heckmatt scale. Ideally, the progression from grade I to IV would represent a linear increase in the extent of muscle pathology. However, from some preliminary comparisons between visual and quantitative grading in 500 patients from our clinic we have seen that Heckmatt grade II muscles have an almost 50-50 chance of being either normal or abnormal, and that there is no clear difference in echogenicity z-score levels between grade III and IV muscles (unpublished observations). This suggest that we may need to rethink the current grading system, for example by using a simpler method that just scores the muscles visually as being either "normal", "uncertain" or "abnormal", or try and construct a more meaningful scale using a Rasch-type method. To further help observers we would need to determine those pitfalls that need to be considered for different muscles, as in knowing for example that tibialis anterior will always be graded Heckmatt grade III in people over 80 years old.

While all this emphasizes that you cannot blindly rely on the results of visual grading and that quantitative evaluation is currently still preferred for optimal sensitivity, even quantified echogenicity measurements have known pitfalls that limit their sensitivity in some situations. First, as the ultrasound beam is attenuated by tissue and more so by more abnormal tissue layers, deeper muscles will show fewer reflections than superficial muscles. A deeply located region of interest is therefore susceptible to a false-negative outcome assessment, appearing blacker than it would be if it were superficial. This makes both quantitated and visual muscle ultrasound unsuitable for the assessment of these deeper muscle layers. Also, complete fatty degeneration of diseased muscle will result in a decrease in grayscale level, and once again a "normal" echogenicity similar to subcutaneous fat (Figure 5G) even though the muscle architecture is clearly disturbed.

All the current pitfalls in image evaluation underscore the need for either a more precise method to visually establish and track abnormality in different muscles, different patients and different disorders, or, alternatively, the need to try and find a more straightforward method to quantify echogenicity and other muscle ultrasound parameters.

5 | SETTING-RELATED CHALLENGES

Even when muscle ultrasound images and videos are perfectly created and interpreted, there are still hurdles to overcome before the technique can be widely implemented. Some of those issues that seem most prominent at this time are discussed below.

5.1 | Ultrasound device availability

Practitioners who initiate neuromuscular ultrasound can find themselves limited by the lack of appropriate ultrasound equipment in their department. In many hospitals, diagnostic general ultrasound is mainly performed by radiologists and in emergency rooms, and time-sharing arrangements for the equipment often poses logistic and financial challenges. As neuromuscular ultrasound is still a relatively new diagnostic technique, the need for specific equipment may not be recognized yet by local hospital administrators. The advent of combined EMG-ultrasound equipment and the development of dedicated or handheld, affordable high-frequency ultrasound devices for muscle
ultrasound could potentially facilitate the decision to start using the technique in the clinical neurophysiology laboratory.

5.2 | Training and certification

Specific hands-on training in muscle ultrasound is not yet a part of standard clinical neurophysiology training programs. An international consensus-based guideline statement for neuromuscular ultrasound training was recently published\(^\text{[72]}\) to help establish a core curriculum, in which visual muscle ultrasound recognition and grading is considered one of the basic techniques. For those seeking to learn, the American Association of Neuromuscular and Electrodagnostic Medicine (AANEM) and International Society of Peripheral Neurophysiologic Imaging (ISPNI) organize yearly courses and conference meetings on nerve and muscle ultrasound, and some specialized centers offer training possibilities for visiting clinicians. There is also much variation in certification requirements across countries. In the United States, the AANEM has put out position statements on how to qualify for practice and how to uniformly conduct and report research in the field,\(^\text{[73,74]}\) and other regional societies such as in Germany have published similar initiatives. A recent development that is to be encouraged, is the incorporation of neuromuscular ultrasound in clinical guidelines, although to our knowledge there is not yet a medical guideline that uses muscle ultrasound specifically.\(^\text{[75,76]}\)

5.3 | Regulatory challenges for muscle ultrasound in trial use

Evidence is accumulating that muscle ultrasound is a good biomarker and can be a surrogate for muscle biopsy in muscular dystrophy treatment trials. As it is patient-friendly and easily repeated, it would be an ideal follow-up tool to track histologic changes over time during the course of a trial. This is highlighted in several studies on DMD, spinal muscular atrophy and FSHD in which muscle ultrasound has been shown to reliably quantify disease progression over time.\(^\text{[73,77-79]}\)

Currently though, there is no regulatory approval for the use of muscle ultrasound (nor for muscle MRI) as a primary outcome in this setting. Reminiscent of a catch-22, a successful treatment trial using muscle ultrasound as a biomarker would be needed to show its clinical relevance to obtain regulatory approval; but as the technique is currently not approved by for example the U.S. Food and Drug Administration, study groups cannot include it as such in their trials. Fortunately, some groups are now incorporating muscle ultrasound as one of the auxiliary outcome measures or for targeted muscle biopsy guidance in their trials, for example in FSHD (ClinicalTrials.gov Identifier: NCT04003974; ultrasound in one participating center) and centronuclear myopathy (ClinicalTrials.gov Identifier: NCT04033159). This will provide much needed further information on the specific value and correlation with other clinical measures and biomarkers.

6 | CONCLUSION

Muscle ultrasound is a valuable addition to the neuromuscular toolkit in both the clinic and in a research setting. As the interest in the technique is increasing, this is the time to ramp up our efforts and tackle the current challenges that hamper its widespread implementation and its benefits for the neuromuscular community. Providing guidelines for the next 5-10 years, we hope this review will encourage colleagues to do so and help unleash muscle ultrasound’s full potential.

ACKNOWLEDGEMENTS

The authors are grateful for the input of dr. Christiaan Saris regarding inflammatory myopathies.

CONFLICT OF INTEREST

Dr. Wijntjes has no conflict of interest to disclose. Dr. van Alfen provides muscle ultrasound consultancy for Dynacure; payment goes to her employer.

ETHICAL PUBLICATION STATEMENT

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

REFERENCES

1. Rahmani N, Mohseni-Bandpei MA, Vameghi R, Salavati M, Abdollahi I. Application of ultrasonography in the assessment of skeletal muscles in children with and without neuromuscular disorders: a systematic review. Ultrasound Med Biol. 2015;41:2275-2283.
2. Mah JK, van Alfen N. Neuromuscular ultrasound: clinical applications and diagnostic values. Can J Neurol Sci/J Can Des Sci Neurol. 2016;45:605-619.
3. van Alfen N, Gijsbertse K, de Korte CL. How useful is muscle ultrasound in the diagnostic workup of neuromuscular diseases? Curr Opin Neurol. 2019;31:568-574.
4. Pillen S, Arts IMP, Zwarts MJ. Muscle ultrasound in neuromuscular disorders. Muscle Nerve. 2008;37:679-693.
5. Pillen S, van Alfen N, Zwarts MJ. Muscle ultrasound: a grown-up technique for children with neuromuscular disorders. Muscle Nerve. 2008;38:1213-1214.
6. Heckmatt JZ, Leeman S, Dubowitz V. Ultrasound imaging in the diagnosis of muscle disease. J Pediatr. 1982;101:656-660.
7. Pillen S, van Keimpema M, Nievelstein RAJ, Verrips A, van Kuijssbergen-Raijmann W, Zwarts MJ. Skeletal muscle ultrasound: Visual versus quantitative evaluation. Ultrasound Med Biol. 2006;32:1315-1321.
8. Brockmann K, Becker P, Schreiber G, Neubert K, Brunner E, Bönnemann C. Sensitivity and specificity of qualitative muscle ultrasound in assessment of suspected neuromuscular disease in childhood. Neuromuscul Disord. 2007;17:517-523.
9. Pillen S, Morava E, Van Keimpema M, et al. Skeletal muscle ultrasonography in children with a dysfunction in the oxidative phosphorylation system. Neuropediatrics. 2006;37:142-147.
10. Adler RS, Garofalo G. Ultrasound in the evaluation of the inflammatory myopathies. Inflamm Myopathies. 2009;11:147-164.
11. Lacourpaille L, Gross R, Hug F, et al. Effects of Duchenne muscular dystrophy on muscle stiffness and response to electrically-induced muscle contraction: a 12-month follow-up. Neuromuscul Disord. 2017;27:214-220.
